Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation

Ingrid E. Christophersen*, MD, PhD,[1–3] Michiel Rienstra*, MD, PhD,[4] Carolina Roselli, MSc*,[1,5,6] Xiaoyan Yin*, PhD,[7,8] Bastiaan Geelhoed*, PhD,[4] John Barnard, PhD,[9] Honghuang Lin, PhD,[7,8] Dan E. Arking, PhD,[10], Albert V. Smith, PhD,[11,12] Christine M. Albert, MD, MPH,[13] Mark Chaffin, MSc[1], Nathan R. Tucker, PhD,[1,2] Molong Li, MD,[2] Derek Klarin, MD,[1] Nathan A Bihlmeyer, BS,[14] Siew-Kee Low, PhD,[15] Peter E. Weeke, MD, PhD,[16,17] Martina Müller-Nurasyid, PhD,[5,18,19] J. Gustav Smith, MD, PhD,[1,20] Jennifer A. Brody, BA,[21] Maartje N. Niemeijer MD,[22] Marcus Dörr, MD,[23,24] Stella Trompet, PhD,[25] Jennifer Huffman, PhD,[26] Stefan Gustafsson, PhD,[27] Claudia Schurman, PhD,[28,29] Marcus E. Kleber, PhD,[30] Leo-Pekka Lyytikäinen, MD,[31] Ilkka Seppälä, MD,[31] Rainer Malik, PhD,[32] Andrea R. V. R. Horimoto, PhD,[33] Marco Perez, MD,[34] Juha Sinisalo, MD, PhD,[35] Stefanie Aeschbacher, MSc,[36,37] Sébastien Thériault, MD, MSc,[38,39] Jie Yao, MS,[40] Farid Radmanesh, MD, MPH,[1,41] Stefan Weiss, PhD,[24,42] Alexander Teumer, PhD,[24,43] Seung Hoan Choi, PhD,[1] Lu-Chen Weng, PhD[1,2] Sebastian Clauss, MD,[2,18] Rajat Deo, MD, MTR,[44] Daniel J. Rader, MD,[44] Svati Shah, MD, MHS,[45] Albert Sun, MD,[45] Jemma C. Hopewell, PhD,[46] Stephanie Debette, MD, PhD,[47–50] Ganesh Chauhan, PhD,[47,48] Qiong Yang, PhD,[51] Bradford B. Worrall, MD, MSc,[52], Guillaume Paré, MD, MSc,[38,39] Yoichiro Kamatani, MD, PhD,[15] Yanick P. Hagemeijer, MSc,[4] Niek Verweij, PhD,[4] Joylene E. Siland, MSc,[4] Michiaki Kubo, MD, PhD,[53] Jonathan D. Smith, PhD,[9] David R. Van Wagoner, PhD,[9] Joshua C. Bis, PhD,[21] Siegfried Perz, MSc,[54] Bruce M. Psaty, MD, PhD,[21,55–57] Paul M. Ridker, MD, MPH,[13] Jared W. Magnani, MD, MSc,[7,58] Tamara B. Harris, MD, MS,[59] Lenore J. Launer, PhD,[59] M. Benjamin Shoemaker, MD, MSCI,[16] Sandosh Padmanabhan, MD,[60] Jeffrey Haessler, MS,[61] Traci M. Bartz, MS,[62] Melanie Waldenberger, PhD,[19,54,63] Peter Lichtner, PhD,[64] Marina Arendt, MSc,[65] Jose E. Krieger, MD, PhD,[33] Mika Kähönen, MD, PhD,[66] Lorenz Risch, MD, MPH,[67] Alfredo J. Mansur, MD, PhD,[68] Annette Peters, PhD,[19,54,69] Blair H. Smith, MD,[70] Lars Lind, MD, PhD,[71] Stuart A. Scott, PhD,[72] Yingchang Lu, MD, PHD,[28,29] Erwin B. Bottinger, MD,[28,73] Jussi Hernesniemi, MD, PhD,[31,74] Cecilia M. Lindgren, PhD,[75] Jorge Wong, MD,[76] Jie Huang, MD, MPH,[77] Markku Eskola, MD, PhD,[74] Andrew P. Morris, PhD,[75,78] Ian Ford, PhD,[79] Alex P. Reiner, MD, MSc,[61,80] Graciela Delgado, Msc,[30] Lin Y. Chen, MD, MS,[81] Yii-Der Ida Chen, PhD,[40] Roopinder K. Sandhu, MD, MPH,[82] Man Li, PhD,[83,84] Eric Boerwinkle, PhD,[85] Lewin Eisele, MD,[65] Lars Lannfelt, MD, PhD,[86] Natalia Rost, MD, MPH, FAAN,[1,87] Christopher D. Anderson, MD, MMSc,[1,41] Kent D. Taylor, PhD,[40] Archie Campbell, MA,[88] Patrik K. Magnusson, PhD,[89] David Porteous, PhD,[88] Lynne J. Hocking, PhD,[90] Efthymia Vlachopoulou, PhD,[91] Nancy L. Pedersen, MA, PhD,[89] Kjell Nikus, MD, PhD,[74] Marju Orho-Melander, PhD,[92] Anders Hamsten, MD, PhD,[93] Jan Heeringa, MD, PhD,[22] Joshua C. Denny, MD,[16] Jennifer Kriebel, PhD,[54,63,69] Dawood Darbar, MD,[94] Christopher Newton-Cheh, MD, MPH,[1,2] Christian Shaffer, BS,[16] Peter W. Macfarlane, PhD, DSc,[95] Stefanie Heilmann, PhD,[96,97] Peter Almgren, MSc,[92] Paul L. Huang, MD, PhD,[2] Nona Sotoodehnia, MD, MPH,[98] Elsayed Z. Soliman, MD, MSc, MS,[99] Andre G. Uitterlinden, PhD,[100] PhD, Albert Hofman, MD, PhD,[22] Oscar H. Franco, MD, PhD,[22] Uwe Völker, PhD,[24,42] Karl-Heinz Jöckel, PhD,[65] Moritz F. Sinner, MD, MPH,[18,19] Henry J. Lin, MD,[40] Xiuqing Guo, PhD,[40] METASTROKE Consortium of the ISGC, Neurology Working Group of the CHARGE Consortium, Martin Dichgans, MD,[32,101,102], Erik Ingelsson, MD, PhD,[27,103] Charles Kooperberg, PhD,[61] Olle Melander, MD, PhD,[104] Ruth J. F. Loos, PhD,[28,29,105] Jari Laurikka, MD, PhD,[106] David Conen, MD, MPH,[36–38] Jonathan Rosand, MD, MSc,[1,41] Pim van der Harst, MD, PhD,[4] Marja-Liisa Lokki, PhD,[91] Sekar Kathiresan, MD,[1] Alexandre Pereira, MD, PhD,[107] J. Wouter Jukema, MD, PhD,[25,108,109] Caroline Hayward, PhD,[26] Jerome I. Rotter, MD,[110] Winfried März, MD,[111] Terho Lehtimäki, MD, PhD,[31] Bruno H. Stricker, MD, PhD,[112] Mina K. Chung, MD,

PhD,[9] Stephan B. Felix, MD,[23,24] Vilmundur Gudnason, MD, PhD,[11,12] Alvaro Alonso, MD, PhD,[113] Dan M. Roden, MD,[16] Stefan Kääb, MD, PhD,[18,19] Daniel I. Chasman, PhD,[1,114] Susan R. Heckbert, MD, PhD,[55,56] Emelia J. Benjamin†, MD, ScM,[7,58,115] Toshihiro Tanaka†, MD, PhD,[116,117] Kathryn L. Lunetta†, PhD,[7,8] Steven A. Lubitz†, MD, MPH,[1,2,118] Patrick T. Ellinor†, MD, PhD,[1,2,118] for the AFGen Consortium

*Contributed equally
†Jointly supervised the work

Affiliations

1. Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA.
2. Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA.
3. Department of Medical Research, Bærum Hospital, Vestre Viken Hospital Trust, Norway.
4. Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.
5. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
6. Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.
7. NHLBI and Boston University's Framingham Heart Study, Framingham, MA, USA.
8. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.
9. Departments of Cardiovascular Medicine, Cellular and Molecular Medicine, Molecular Cardiology, and Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, USA.
10. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
11. Icelandic Heart Association, Kopavogur, Iceland.
12. Faculty of Medicine, University of Iceland, Reykavik, Iceland.
13. Divisions of Preventive and Cardiovascular Medicine, Brigham and Women's Hospital & Harvard Medical School, Boston, MA, USA.
14. Predoctoral Training Program in Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
15. Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
16. Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA.
17. The Heart Centre, Department of Cardiology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark.
18. Department of Medicine I, University Hospital Munich, Ludwig-Maximilians-University, Munich, Germany.
19. DZHK (German Centre for Cardiovascular Research), partner site: Munich Heart Alliance, Munich, Germany.
20. Molecular Epidemiology and Cardiology, Clinical Sciences, Lund University, Lund, Sweden.
21. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA.

22. Department of Epidemiology, Erasmus University Medical Center Rotterdam, Rotterdam, the Netherlands.
23. Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany.
24. DZHK (German Centre for Cardiovascular Research), partner site: Greifswald, Germany.
25. Department of Cardiology, Leiden University Medical Center, The Netherlands.
26. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, UK.
27. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.
28. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
29. The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
30. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Germany.
31. Department of Clinical Chemistry, Fimlab Laboratories and University of Tampere School of Medicine, Tampere, Finland.
32. Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians University, München, Germany.
33. Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil.
34. Stanford University, Stanford, CA, USA.
35. Heart and Lung Center HUS, Helsinki University Central Hospital, Helsinki, Finland.
36. University Hospital Basel, Switzerland.
37. Cardiovascular Research Institute Basel, Switzerland.
38. Population Health Research Institute, Hamilton, Canada.
39. Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Canada.
40. Institute for Translational Genomics and Population Sciences, Department of Pediatrics, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
41. Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA.
42. Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany.
43. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany.
44. Division of Cardiovascular Medicine, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA.
45. Division of Cardiology, Department of Medicine, Duke University School of Medicine, Durham, NC, USA.
46. CTSU - Nuffield Department of Population Health, University of Oxford, Oxford, UK.
47. Inserm Center U1219 (Bordeaux Population Health Centre), Bordeaux, France.
48. University of Bordeaux, Bordeaux, France.
49. Department of Neurology, Bordeaux University Hospital, Bordeaux, France.
50. Department of Neurology, Boston University School of Medicine, Boston, MA, USA.
51. Biostatistics Department, School of Public Health, Boston University, Boston, MA, USA.
52. University of Virginia Health System, Departments of Neurology and Public Health Science, Charlottesville, VA, USA.

53. RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
54. Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
55. Department of Epidemiology and Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA.
56. Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA.
57. Department of Health Services, University of Washington, Seattle, WA, USA.
58. Department of Medicine, Boston University School of Medicine, Boston, MA, USA.
59. Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, MD, USA.
60. Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, UK.
61. Fred Hutchinson Cancer Research Center, University of Washington, Seattle, WA, USA.
62. Cardiovascular Health Research Unit, Departments of Medicine and Biostatistics, University of Washington, Seattle, WA, USA.
63. Research unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
64. Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
65. Institute for Medical Informatics, Biometry, and Epidemiology, University Hospital, University Duisburg-Essen, Germany.
66. Department of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
67. University Institute of Clinical Chemistry, University of Bern, Switzerland and labormedizinisches zentrum Dr. Risch, Schaan, Liechtenstein.
68. Heart Institute, University of Sao Paulo, Sao Paulo, Brazil.
69. German Center for Diabetes Research, Neuherberg, Germany.
70. Division of Population Health Sciences, University of Dundee, Scotland, UK.
71. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden.
72. Department of Genetics and Genomic Sciences , Icahn School of Medicine at Mount Sinai, New York, NY, USA.
73. Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
74. Department of Cardiology, Heart Hospital, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
75. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
76. Division of Cardiology, Hamilton Health Sciences, McMaster University, Hamilton, Ontario, Canada.
77. Boston VA Research Institute, Inc., Boston, MA, USA.
78. Department of Biostatistics, University of Liverpool, Liverpool, UK.
79. Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK.
80. Department of Epidemiology, University of Washington, Seattle, WA, USA.

81. Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, MN, USA.
82. Division of Cardiology, University of Alberta, Edmonton, Canada.
83. Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA.
84. Division of Nephrology & Hypertension, Internal Medicine, School of Medicine, University of Utah, UT, USA.
85. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA.
86. Department of Public Health and Caring Sciences, Geriatrics, Uppsala University, Uppsala, Sweden.
87. Acute Stroke Services, Massachusetts General Hospital, Boston, MA, USA.
88. Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, UK.
89. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
90. Musculoskeletal Research Programme, Division of Applied Medicine, University of Aberdeen, Aberdeen, UK.
91. Transplantation Laboratory, Medicum, University of Helsinki, Helsinki, Finland.
92. Department of Clinical Sciences, Lund University, Malmö, Sweden.
93. Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.
94. University of Illinois, Chicago, IL, USA.
95. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK.
96. Institute of Human Genetics, University of Bonn, Germany.
97. Department of Genomics, Life & Brain Research Center, University of Bonn, Germany.
98. Cardiovascular Health Research Unit, University of Washington Medical Center, Seattle, WA, USA.
99. Epidemiological Cardiology Research Center (EPICARE), Wake Forest School of Medicine, Winston Salem, NC, USA.
100. Department of Epidemiology and Internal Medicine, Erasmus University Medical Center Rotterdam, the Netherlands.
101. Munich Cluster for Systems Neurology (SyNergy), München, Germany.
102. German Center for Neurodegenerative Diseases (DZNE), Munich, Germany.
103. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA.
104. Department of Internal Medicine, Clinical Sciences, Lund University, Malmö, Sweden.
105. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
106. Department of Cardio-Thoracic Surgery, Heart Hospital, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
107. Laboratory of Genetics and Molecular Biology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil and Department of Genetics, Harvard Medical School, Boston, MA, USA.
108. Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands.
109. Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands.

110. Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
111. Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria and Synlab Academy, Synlab Services GmbH, Mannheim, Germany.
112. Department of Epidemiology and Internal Medicine, Erasmus University Medical Center Rotterdam, the Netherlands and Inspectorate of Health Care, Utrecht, the Netherlands.
113. Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA.
114. Divisions of Preventive Medicine and Genetics, Brigham and Women's Hospital & Harvard Medical School, Boston, MA, USA.
115. Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.
116. Laboratory for Cardiovascular Diseases, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
117. Department of Human Genetics and Disease Diversity, Tokyo Medical and Dental University Graduate School of Medical and Dental Sciences, Tokyo, Japan.
118. Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA.

Word count: 2,229 (excl. online methods, references, and figure legends)

Short title: Novel genetic loci for atrial fibrillation

These authors contributed equally to this work: Ingrid E. Christophersen, Michiel Rienstra, Carolina Roselli, Xiaoyan Yin, Bastiaan Geelhoed.

These authors jointly supervised this work: Emelia J. Benjamin, Toshihiro Tanaka, Kathryn L. Lunetta, Steven A. Lubitz, Patrick T. Ellinor.

Corresponding author:
Patrick T. Ellinor, MD, PhD,
Program in Medical and Population Genetics,
The Broad Institute of MIT and Harvard

Cardiovascular Research Center,
Massachusetts General Hospital;
Boston, MA 02129
T: 617-724-8729
E: ellinor@mgh.harvard.edu

Journal subject codes: Atrial fibrillation, population genetics, genome-wide association studies, gene-expression

Disclosures: Dr. Ellinor is the PI on a grant from Bayer HealthCare to the Broad Institute focused on the genetics and therapeutics of atrial fibrillation. The remaining authors have no disclosures.

Atrial fibrillation affects more than 33 million people worldwide and increases the risk of stroke, heart failure, and death.[1,2] Fourteen genetic loci have been associated with atrial fibrillation in European and Asian ancestry groups.[3–7] To further define the genetic basis of atrial fibrillation, we performed large-scale, multi-racial meta-analyses of common and rare variant association studies. The genome-wide association studies (GWAS) included 18,398 individuals with atrial fibrillation and 91,536 referents; the exome-wide association studies (ExWAS) and rare variant association studies (RVAS) involved 22,806 cases and 132,612 referents. We identified 12 novel genetic loci that exceeded genome-wide significance, implicating genes involved in cardiac electrical and structural remodeling. Our results nearly double the number of known genetic loci for atrial fibrillation, provide insights into the molecular basis of atrial fibrillation, and may facilitate new potential targets for drug discovery.[8]

Atrial fibrillation is a common cardiac arrhythmia that can cause serious complications such as stroke, heart failure, dementia, and death.[1,2] The lifetime risk of atrial fibrillation is one in four[9] and it has been estimated that more than 33 million individuals worldwide are affected.[1] During the last decade, GWAS have identified 13 genetic loci associated with atrial fibrillation in Europeans and one Asian specific atrial fibrillation locus, of which a region near the gene encoding the transcription factor PITX2 has shown the strongest association.[3–7] Recently, genome and exome sequencing studies have identified rare atrial fibrillation-associated mutations in *MYL4*,[10] *MYH6*,[11] *CACNB2,*[12] and *CACNA2D4.*[12] Given the incomplete understanding of the biology of atrial fibrillation and the modestly sized prior genetic association analyses, we sought to identify additional susceptibility loci by increasing the size and diversity of the atrial fibrillation studies.

We therefore investigated both common and rare variants in a large collection of individuals in the Atrial Fibrillation Genetics (AFGen) Consortium, by meta-analyses of GWAS, ExWAS, and RVAS in 33

studies, including 22,806 individuals with atrial fibrillation and 132,612 referents (**Online methods**). **Fig. 1** illustrates our study design and **Supplementary Tables 1 and 2** show baseline characteristics of the study participants.

In a meta-analysis of GWAS in 31 studies, we identified 10 new genetic loci associated with atrial fibrillation ($P < 5x10^{-8}$) at *METTL11B/KIFAP3, ANXA4/GMCL1, CEP68, TTN/TTN-AS1, KCNN2, KLHL3/WNT8A/FAM13B, SLC35F1/PLN, ASAH1/PCM1, SH3PXD2A,* and *KCNJ5* (**Table 1, Figs. 2 and 3, Supplementary Fig. 1, Supplementary Table 3**). The 13 genetic loci previously associated with atrial fibrillation in Europeans were again observed, while one locus previously reported in Asians only, did not reach genome-wide significance in our study (*CUX2*).

In a meta-analysis of ExWAS in 17 studies, we identified two additional novel genetic loci (*SCN10A* and *SOX5,* $P < 1.04x10^{-6}$) as well as one new locus also identified in the GWAS meta-analysis (*SLC35F1/PLN*) (**Table 2, Supplementary Fig. 2 and 3**). Variants at each of these three loci have previously been associated with electrocardiographic traits (**Supplementary Table 3**).

Finally, in an RVAS or burden test of rare variants, one gene, *SH3PXD2A*, reached genome-wide significance. This association was mainly driven by a rare coding variant that is unique to individuals of Asian ancestry (rs202011870, minor allele frequency (MAF) 0.18%, odds ratio (OR) 4.68, 95% confidence interval (CI) 2.97-7.39, $P=3.3x10^{-11}$, **Supplementary Tables 3-5**) and the same locus was significantly associated with atrial fibrillation in the GWAS meta-analysis. Out of the 11 variants in the Asian ancestry burden test, rs149867987 also reached genome-wide significance and had an effect in the same direction as rs202011870. There was no genome-wide significant signal at *SH3PXD2A* in RVAS analyses in individuals of European or African American ancestry.

Ancestry-specific GWAS analysis revealed a significant association between African Americans (641 cases and 4956 referents) with atrial fibrillation and variants on chromosome 4q25 upstream of *PITX2* (rs6843082, OR 1.40, 95% CI 1.24-1.58, P=4.31x10^{-8}, **Supplementary Table 6, Supplementary Fig. 4**). Similarly, the 4q25/*PITX2* region is the most significant locus for atrial fibrillation in individuals of Japanese ancestry (rs2723334, OR 1.94, 95% CI 1.68-2.25, P=8.46x10^{-19}) and European ancestry (rs2129977, OR 1.45, 95% CI 1.41-1.49, P=7.25x10^{-136}), and the lead SNPs in all three ancestry groups are in strong linkage disequilibrium, with an r^2>0.94. Further ancestry-specific meta-analyses did not produce additional robust associations for atrial fibrillation **(Supplementary Results, Supplementary Table 6-7, and Supplementary Figs. 4-6)**. Separate meta-analyses of incident and prevalent atrial fibrillation in Europeans did reveal one additional genome wide signal at chromosome 12p11/*PKP2* that was only present in the prevalent atrial fibrillation analysis (**Supplementary Results, Supplementary Tables 8-9, Supplementary Figs. 7-8**); however, since this locus was not present in the combined analyses it was not pursued further.

We then performed an *in silico* replication of our results using two ethnically distinct studies. First, we replicated the atrial fibrillation associated variants in 8,180 cases and 28,612 referents from the Biobank Japan study (**Online methods, Supplementary Table 10**). The novel atrial fibrillation variant intronic to *CEP68* reached genome-wide significance among Japanese, whereas the atrial fibrillation variants at *KCNN2* and *SOX5* achieved significance when correcting for multiple testing of 33 variants (P<1.5x10^{-3}). The loci at *ASAH1, TTN,* and *METTL11B* reached nominal significance in Japanese (P<0.05). Of note, approximately 10% of the cases in the GWAS discovery analysis and Japanese replication analysis were overlapping (837 cases and 3293 referents). The lack of replication of the remaining loci likely reflects the heterogeneous nature of atrial fibrillation across different ancestries.

Second, we performed replication in 3,366 cases and 139,852 referents of mainly European ancestry in the UK Biobank (**Online methods, Supplementary Table 11**). The atrial fibrillation locus at *SH3PXD2A* reached genome-wide significance in the UK Biobank, whereas the loci *METTL11B, CEP68,* and *KLHL3/WNT8A/FAM13B* were significantly associated when correcting for multiple testing of 31 variants ($P<1.6x10^{-3}$), and the loci at *TTN, ASAH1*, *KCNJ5,* and *SCN10A* reached nominal significance ($P<0.05$). The lack of replication of all of the atrial fibrillation loci is likely caused by reduced statistical power due to decreased sample size in the replication sample (18,398 versus 3,366 atrial fibrillation cases). However, there was a consistent direction of effects for all atrial fibrillation loci in the discovery and replication analyses.

Conditional analyses based on the summary level results of the GWAS meta-analysis were performed to identify multiple, independent signals on each chromosome containing atrial fibrillation loci (**Online Methods**). We confirmed that the two loci *METTL11B/KIFAP3* and *PRRX1,* located ~350 kilobases (kb) apart on chromosome 1, were independent signals, as were the two loci *SH3PXD2A* and *NEURL1,* ~200 kb apart on chromosome 10 (**Supplementary Table 12, Supplementary Fig. 9**).

We found that seven of the known or new atrial fibrillation loci were associated with atrial fibrillation-related phenotypes, such as electrocardiographic traits, left ventricle internal diastolic diameter, and stroke (**Supplementary Table 3 and 13, Supplementary Fig. 10**). Given the close relation between atrial fibrillation and cardioembolic stroke, we then sought to determine whether the novel atrial fibrillation variants were associated with stroke risk. We performed an *in silico* lookup in GWAS data for stroke subtypes from the Neuro-CHARGE and METASTROKE consortia. None of the novel loci for

atrial fibrillation were associated with ischemic stroke, cardioembolic stroke, small, or large vessel disease (**Supplementary Tables 14-15**).

Next, we performed an *in silico* evaluation of the known and newly identified atrial fibrillation associated loci (**Online Methods, Supplementary Results**). We compared the atrial fibrillation loci (n=24) to other trait-associated loci from the NHGRI-EBI GWAS catalog (n=3,381) and matching control loci selected for similar architectural properties (n=9,093). Interestingly, the atrial fibrillation loci were significantly conserved across species, and were also significantly enriched for active enhancers in cardiac tissues as denoted by H3K27ac marks, compared to other trait-associated loci from the NHGRI-EBI GWAS catalog and matching control loci (**Supplementary Fig. 11**). Moreover, the genes at atrial fibrillation loci displayed enrichment for Gene Ontology terms important for cardiac action potential propagation and cardiac contractility compared to the control loci, although this enrichment was not significant when corrected for multiple hypothesis testing (**Supplementary Table 16**).

We also performed expression quantitative trait locus (eQTL) analyses of the atrial fibrillation-associated genetic loci using two additional approaches (**Online Methods**). We identified significant eQTLs for seven of the twelve novel atrial fibrillation associated loci (closest gene;eQTL gene: *METTL11B;KIFAP3, ANXA4;ANXA4/GMCL1/PCYOX1/SNRNP27, CEP68;CEP68, KCNN2;KCNN2, KLHL3;FAM13B/REEP2, ASAH1;ASAH1/PCM1/RP11-806O11.1,* and *KCNJ5*;*KCNJ5/C11orf45*) and eight of the thirteen previously reported atrial fibrillation loci (**Supplementary Tables 17-20, Supplementary Fig. 12**).

In the current work, we have identified 12 novel genetic loci for atrial fibrillation in our large-scale analyses of common, coding, and rare genetic variation for atrial fibrillation (**Supplementary Table 3)**.

When considered together with the known atrial fibrillation loci, the genes at these loci broadly encode ion channels, sarcomeric proteins, and transcription factors that underlie this common arrhythmia. Genes at five of the genetic loci identified encode potassium or sodium channels, including two novel loci at the genes *KCNN2* and *KCNJ5* that are known to be involved in the maintenance of the atrial cardiac action potential. Since the cellular hallmark of atrial fibrillation is shortening of the atrial action potential duration and calcium overload, the *KCNN2* and *KCNN3* genes are particularly interesting. The lead variant at chromosome 5q22 is located intronic to and has a significant eQTL with *KCNN2*, which encodes the calcium dependent potassium channel SK2. The SK2 protein is known to form heteromeric channel complexes with SK3, which is a product of the *KCNN3* gene that is strongly associated with atrial fibrillation in the present and previous atrial fibrillation GWAS meta-analyses.[5,6]

Similarly, *KCNJ5* encodes the potassium channel Kir3.4 or GIRK4 that is known to form heteromeres with Kir3.1/GIRK1/*KCNJ3* and assemble to form the inwardly rectifying, I_{KAch} channel complex. The I_{KAch} complex is regulated by G protein signaling, is well-known to regulate the membrane potential in the sinoatrial node and atria, and has been considered as a therapeutic target for atrial fibrillation.

Interestingly, the gene identified in our rare and common variant analyses, *SH3PXD2A*, is expressed in human atria and ventricles and encodes TKS5, a tyrosine kinase substrate. The rare variant association was largely driven by the variant rs202011870, which results in a leucine to arginine substitution at position 396. TKS5 has been shown to be important in determining the invasiveness of cancer cells[13] and has been suggested to mediate the neurotoxic effect of beta-amyloid in Alzheimer disease in association with the matrix metalloproteinase gene *ADAM12*.[14] Developmentally, *SH3PXD2A* is important for neural crest migration; homozygous knockout in mice result in complete cleft in the secondary palate and

neonatal death;[15] however, the relation between *SH3PXD2A* and atrial fibrillation is unclear and as with any rare variant association, replication in a large, independent dataset will ultimately be required.

Finally, we found that the atrial fibrillation loci have significant conservation across species, and are enriched for active enhancers in cardiac tissues, compared to other GWAS or control loci. Since many of the identified atrial fibrillation loci include genes that encode transcription factors (*PITX2*, *ZFHX3*, *PRRX1*, *SOX5*, and *TBX5*), we hypothesize that these loci may be more conserved, because they may underlie a canonical program for left atrial and/or pulmonary venous development.

While the strengths of our study include the large sample sizes, analyses of common and rare genetic variation, and the inclusion of different races and ethnicities, our study was subject to some limitations. Specifically, it is important to note that the estimates of variance explained by genetic variation can be challenging for qualitative traits such as atrial fibrillation, particularly given the marked variability in prevalence of the disease according to age. Thus, as with GWAS for other common conditions, we anticipate that the newly described loci for atrial fibrillation would only explain a small portion of the variance of atrial fibrillation.

In conclusion, we have nearly doubled the number of known genetic loci associated with atrial fibrillation through meta-analysis of more than 22,000 individuals with atrial fibrillation. We have identified a series of novel atrial fibrillation-associated variants, which lie proximal to genes involved in atrial electrical and mechanical function. Our results will facilitate downstream research establishing the mechanistic links between identified genetic loci and atrial fibrillation pathogenesis, potentially aiding in the discovery of new therapeutic targets for the treatment of atrial fibrillation.[8]

Code availability

The computer code that support the results of the present study are available from the corresponding author upon request.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

A full list of acknowledgments appears in the **Supplementary Note**.

Author Contributions

I.E.C., C.R., X.Y., T.T., K.L.L., E.B.J, S.A.L., M.R., B.G., P.T.E. wrote and edited the manuscript. All authors contributed to and discussed the results, and commented on the manuscript. GWAS and ExWAS analyses: A.V.S, N.A.B., M.M-N., I.S., C.S., P.E.W., S.A., S.T., J.A.B., J.C.B., H.L., J.H., J.Y., X.G., F.R., M.N.N., D.E.A., G.P., S-K.L., Y.K., M.K., A.C.P., A.R.H., J.S., L-P.L., M.A., M.E.K., J.G.S., R.M., S.G., S.T., M.D., S.W., J.W., D.I.C., M.V.P., Q.Y., T.B.H., M.F.S., J.S., D.v.W., M.K. Individual dataset quality control and GWAS and ExWAS meta-analyses: I.E.C., K.L.L., C.R., X.Y., M.R., B.G., Y.P.H., N.V., J.E.S. Replication in METASTROKE and Neuro-CHARGE: Q.Y., J.H., S.D., G.C., B.B.W. Replication in UK Biobank: S.K., D.K., C.N-C. Replication in Biobank Japan: S-K.L., Y.K., M.K., T.T. Replication in African American population: R.D., D.J.R., S.S., A.S. CCAF eQTL analyses: J.B., M.K.C., D.v.W., J.D.S. Functional annotation: I.E.C., S.H.C., L-C.W., M.L., C.R., M.C., N.R.T., S.C. Pathway analyses: H.L.

Competing Financial Interests Statement

Dr. Ellinor is the PI on a grant from Bayer HealthCare to the Broad Institute focused on the genetics and therapeutics of atrial fibrillation. The remaining authors have no disclosures.

References

1. Chugh, S. S. *et al.* Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* **129,** 837–47 (2014).
2. January, C. T. *et al.* 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J. Am. Coll. Cardiol.* **64,** 2246–2280 (2014).
3. Gudbjartsson, D. F. *et al.* Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* **448,** 353–357 (2007).
4. Benjamin, E. J. *et al.* Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat. Genet.* **41,** 879–81 (2009).
5. Ellinor, P. T. *et al.* Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat. Genet.* **42,** 240–4 (2010).
6. Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat. Genet.* **44,** 670–5 (2012).
7. Sinner, M. F. *et al.* Integrating genetic, transcriptional, and functional analyses to identify 5 novel genes for atrial fibrillation. *Circulation* **130,** 1225–35 (2014).
8. Nelson, M. R. *et al.* The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47,** 856–860 (2015).
9. Lloyd-Jones, D. M. *et al.* Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. *Circulation* **110,** 1042–1046 (2004).
10. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47,** 435–444 (2015).
11. Holm, H. *et al.* A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat.*

Genet. **43,** 316–20 (2011).

12. Weeke, P. *et al.* Exome sequencing implicates an increased burden of rare potassium channel variants in the risk of drug-induced long QT interval syndrome. *J. Am. Coll. Cardiol.* **63,** 1430–7 (2014).

13. Seals, D. F. *et al.* The adaptor protein Tks5/Fish is required for podosome formation and function, and for the protease-driven invasion of cancer cells. *Cancer Cell* **7,** 155–65 (2005).

14. Laumet, G. *et al.* A study of the association between the ADAM12 and SH3PXD2A (SH3MD1) genes and Alzheimer's disease. *Neurosci. Lett.* **468,** 1–2 (2010).

15. Cejudo-Martin, P. *et al.* Genetic disruption of the sh3pxd2a gene reveals an essential role in mouse development and the existence of a novel isoform of tks5. *PLoS One* **9,** e107674 (2014).

16. Pruim, R. J. *et al.* LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26,** 2336–2337 (2010).

Figure legends

Figure 1. Study flow-chart.

Overview of the approach employed for genome-wide and exome-wide association analyses.

Figure 2. Manhattan plot of the combined ancestry GWAS meta-analyses.

Manhattan plot showing novel (red) and replicated (blue) genetic loci associated with atrial fibrillation in the combined ancestry GWAS meta-analysis. The dotted line represents the threshold of statistical significance ($5x10^{-8}$). The gene names represent the gene in closest proximity to the most significant variant at each locus. There is a break in the Y-axis to increase the resolution of the genetic loci near the genome-wide significance threshold.

Figure 3. Regional plots from combined ancestry GWAS meta-analysis.

The most significant variant at each locus is plotted (purple, diamond-shaped) and identified with rsID. Each dot in the plots represent a single variant present in our results and the color of the dot indicates the degree of linkage disequilibrium with the most significant variant, as shown on the top left color chart on each panel. The lower part of each panel shows the locations of genes at the respective loci. r^2, degree of linkage disequilibrium; chr, chromosome; Mb, megabases; cM, centiMorgan. Regional plots were created using LocusZoom.[16]

Table 1. Results from combined ancestry GWAS meta-analysis

rsID	Chr	Gene(s)	Location relative to gene	Risk allele/ reference allele	Risk allele frequency, %	OR	95% CI	P-value	Mean imputation quality
Novel associations									
rs72700118	1q24	*METTL11B/KIFAP3*	Intergenic	A/C	12	1.14	1.10-1.19	2.60×10^{-11}	0.959
rs3771537	2p13	***ANXA4****/GMCL1*	Intronic	A/C	53	1.09	1.06-1.12	7.92×10^{-12}	0.987
rs2540949	2p14	***CEP68***	Intronic	A/T	61	1.08	1.06-1.11	2.93×10^{-10}	0.991
rs2288327	2q31	***TTN/TTN-AS1***	Intronic	G/A	20	1.09	1.06-1.13	2.05×10^{-8}	0.994
rs337711	5q22	***KCNN2***	Intronic	T/C	39	1.07	1.05-1.10	2.93×10^{-8}	0.995
rs2967791	5q31	***KLHL3****/WNT8A/FAM13B*	Intronic	T/C	54	1.07	1.05-1.10	2.73×10^{-8}	0.961
rs4946333	6q22	***SLC35F1****/PLN*	Intronic	G/A	50	1.08	1.05-1.10	1.89×10^{-9}	0.995
rs7508	8p22	***ASAH1****/PCM1*	3'UTR	A/G	72	1.09	1.06-1.12	5.16×10^{-10}	0.977
rs35176054	10q24	***SH3PXD2A***	Intronic	A/T	13	1.14	1.10-1.18	8.63×10^{-12}	0.939
rs75190942	11q24	***KCNJ5***	Intronic	A/C	8	1.17	1.11-1.24	1.59×10^{-8}	0.744
Previously known associations									
rs11264280	1q21	*KCNN3*	Intergenic	T/C	31	1.12	1.09-1.15	6.41×10^{-17}	0.942
rs520525	1q24	***PRRX1***	Intronic	A/G	71	1.12	1.09-1.15	6.39×10^{-16}	0.955
rs11718898	3p25	***CAND2***	Exonic	C/T	65	1.08	1.05-1.10	4.68×10^{-8}	0.969
rs6843082	4q25	*PITX2*	Intergenic	G/A	25	1.45	1.41-1.49	3.41×10^{-155}	0.989
rs12664873	6q22	*GJA1*	Intergenic	T/G	70	1.08	1.05-1.11	1.19×10^{-8}	0.968
rs1997572	7q31	***CAV1/2***	Intronic	G/A	59	1.10	1.08-1.13	6.64×10^{-15}	0.988
rs7026071	9q22	***C9orf3***	Intronic	T/C	40	1.09	1.07-1.12	1.31×10^{-12}	0.970
rs7915134	10q22	*SYNPO2L*	Intergenic	C/T	85	1.12	1.08-1.16	1.68×10^{-10}	0.975
rs11598047	10q24	***NEURL1***	Intronic	G/A	16	1.18	1.14-1.21	1.67×10^{-22}	0.971
rs883079	12q24	***TBX5***	3'UTR	T/C	70	1.11	1.09-1.14	1.80×10^{-15}	0.991
rs1152591	14q23	***SYNE2***	Intronic	A/G	46	1.09	1.06-1.11	1.04×10^{-10}	0.960
rs74022964	15q24	*HCN4*	Intergenic	T/C	17	1.12	1.08-1.15	2.37×10^{-11}	0.970
rs2106261	16q22	***ZFHX3***	Intronic	T/C	19	1.20	1.17-1.24	8.18×10^{-32}	0.973

The most significant variant at each genetic locus associated with atrial fibrillation is listed. Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed. Chr, chromosome; CI, confidence interval; OR, odds ratio.

Table 2. Results from combined ancestry ExWAS meta-analysis

rsID	Chr	Gene(s)	Location relative to gene	Risk allele/ reference allele	Risk allele frequency, %	OR	95% CI	P-value
Novel associations								
rs6800541	3p22	***SCN10A***	Intronic	T/C	61	1.08	1.05-1.12	8.79×10^{-7}
rs89107	6q22	***SLC35F1****/PLN*	Intronic	G/A	58	1.07	1.04-1.10	9.51×10^{-7}
rs11047543	12p12	*SOX5*	Intergenic	G/A	86	1.14	1.10-1.19	2.47×10^{-12}
Previously known associations								
rs13376333	1q21	***KCNN3***	Intronic	T/C	23	1.13	1.09-1.16	1.46×10^{-12}
rs17042171	4q25	*PITX2*	Intergenic	A/C	21	1.64	1.59-1.69	8.31×10^{-227}
rs3807989	7q31	***CAV1***	Intronic	G/A	58	1.09	1.06-1.12	6.52×10^{-8}
rs60632610	10q22	***SYNPO2L***	Exonic; nonsyn	C/T	85	1.12	1.08-1.15	1.54×10^{-10}
rs10151658	14q23	***SYNE2***	Exonic; nonsyn	C/A	49	1.07	1.04-1.09	5.16×10^{-7}
rs2106261	16q22	***ZFHX3***	Intronic	A/G	17	1.21	1.16-1.26	4.00×10^{-19}

The most significant variant at each genetic locus associated with atrial fibrillation is listed. Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed. Chr, chromosome; CI, confidence interval; OR, odds ratio; nonsyn, nonsynonymous.

Online METHODS

Study population

The Atrial Fibrillation Genetics Consortium (AFGen) is a collaboration between multiple studies with the aim of investigating the genetic causes of atrial fibrillation. In this study, we included 33 studies from AFGen, of which 31 participated in the GWAS meta-analysis, whereas 17 studies were part of the exome chip analyses. **Supplementary Table 21** shows per study overlap of samples between the GWAS and exome chip analyses. The majority of the participants were of European ancestry (15,993 cases, 113,719 referents). We also included studies with African-American (3 studies; 641 cases, 4956 referents), Japanese (1 study; 837 cases, 2456 referents), Hispanic (1 study; 277 cases, 3081 referents), and Brazilian (1 study; 187 cases, 550 referents) ancestry (Supplementary Table 1). The ExWAS and RVAS involved 22,806 cases and 132,612 referents of European (13,496 cases, 96,273 referents), African American (681 cases, 4,871 referents), and Asian (8,180 cases, 28,612 referents) ethnicities (**Supplementary Table 2**). Overall, adjudication of atrial fibrillation included either documented atrial fibrillation on an electrocardiogram and/or one in-patient or two out-patient diagnoses of atrial fibrillation. Referents were free of atrial fibrillation. All participating studies had obtained informed consent from all cases and referents and had obtained approval from their respective ethics committees or institutional review boards.

GWAS meta-analyses

Each study performed genotyping and imputation to the 1000 Genomes Project Phase 1 reference panel (March 2012 release). Detailed methods for each study are described in the **Supplementary Note** and in **Supplementary Table 22**. Cox proportional hazards models were used for incident data with time-to-event from study enrollment. Logistic regression models were used for

prevalent and case-control data. Models were adjusted for age and sex if available, and if appropriate, for principal components of the genotype matrix to control for population stratification. For studies with prevalent cases at time of enrollment (or blood draw) and incident cases identified during follow up, two analyses were performed: 1) Prevalent analysis at baseline/blood draw: all individuals who were diagnosed with atrial fibrillation prior to baseline were defined as cases, and all individuals who were not diagnosed with AF prior to baseline were defined as referents in a logistic regression analysis (future cases were controls in this analysis); 2) Incident analysis looking forward from baseline: prevalent cases were excluded and time-to-atrial fibrillation diagnosis was analyzed, using Cox proportional hazards models, with censoring at last follow-up. The two analyses are approximately independent, because they consider different periods of risk, as described by Benjamin et al.[1]

Pre- and post-GWAS filtering was performed according to predefined quality control filters (**Supplementary Table 23**). Briefly, variants with MAF <1%, imputation quality <0.3 (IMPUTE), or that were present in <2 studies were excluded.

We meta-analyzed summary level GWAS results using an inverse variance-weighted fixed-effects model with METAL software.[2] For the combined ancestry GWAS meta-analysis, we tested 11,795,432 variants. The traditional Bonferroni correction for number of variants tested is often regarded as too conservative, because the tests are not independent due to LD. Thus, we chose the most widely used and accepted significance threshold for GWAS in our GWAS meta-analyses.[3–6] Variants that reached a genome-wide P-value $<5x10^{-8}$ were considered statistically significant. Meta-analyses were also performed separately for each ethnicity group and for incident and prevalent atrial fibrillation to identify potentially differential associations and effects.

ExWAS and rare variant meta-analyses

Each study performed exome variant genotyping and association analyses locally, using a logistic model that combined incident and prevalent cases and referents (**Supplementary Table 24**). Individual variants that passed quality control filters and were present in at least 2 studies with average MAF≥0.5% (**Supplementary Table 23**), were meta-analyzed using the score test implemented in the seqMeta package of R statistical software.[7] For the combined ancestry ExWAS meta-analysis, we tested 48,133 variants and used a significance level of 1.04×10^{-6}, which is approximately a Bonferroni adjustment of 0.05/48,133. For MAF > 0.5%, we had approximately 80% power to detect variants with a multiplicative genotype relative risk of 1.4. RVAS was performed on rare variants from the exome chip array using SKAT[8] and burden tests with three approaches: 1) all non-synonymous and splice site variants, 2) non-synonymous variants annotated as possibly damaging, and 3) loss-of-function variants only. For each gene-based test we excluded variants with MAF >5% and excluded genes with cumulative MAF <0.05%.

Approximate joint and conditional analysis

To identify independent variants within the 12 significant genetic loci, we performed an approximate joint and conditional association analysis implemented in the software GCTA[9] using summary level statistics from the meta-analysis. We used a stepwise procedure for detecting additional independent variants with a European ancestry reference panel from the Framingham Heart Study (n=2764 unrelated individuals).

Functional annotation

Functional element enrichment: Loci were defined as regions encompassing variants that were in linkage disequilibrium with the query variant (r^2>0.8 in CEU population) and that were no greater

than 500 kb from the query variant. Loci had to encompass at least 5 kb both upstream and downstream of the query variant. Overlapping loci were merged. The GWAS control loci were calculated from unique variants from the NHGRI-EBI GWAS catalog (as of May 31, 2016) that had a P-value $<5x10^{-8}$. The 1000 Genomes control loci were calculated using 24,000 matched variants based on MAF, gene density, distance to nearest gene, and number of nearby variants in linkage disequilibrium determined by the SNPsnap tool.[10] The SNPsnap matched variants were calculated using the European population and an r2 cutoff of 0.8, but otherwise default parameters. Each locus in each experimental set was intersected with various markers for functional elements to determine the median percent overlap of each experimental set. The markers included phastCons 46-way primate and mammalian conserved elements, Roadmap Epigenome H3K27ac gapped peaks, and ENCODE DNaseHS sites. Statistical significance was calculated by one-tailed bootstrapping for enrichment with 1,000 random sub-samplings of each control set.

Gene ontology analysis of atrial fibrillation loci: RefSeq genes that overlapped atrial fibrillation-associated loci as well as genes that overlapped the GWAS catalog control loci and the 1000 Genomes matched control loci were used for gene ontology enrichment analysis. The genes that overlapped the control loci were used as two separate background sets. Enrichment calculations were provided by the GOrilla tool.[11]

In silico database interrogation: All statistically significant variants and genes from GWAS and RVAS analyses were selected for an in silico assessment through lookups in the following databases: The Gene Tissue Expression database (GTEx),[12] RegulomeDB,[13] HaploREG,[14] GeneCards (www.genecards.org/), dbSNP.[15] From the GTEx search, we report statistically significant eQTLs in cardiac and skeletal muscle tissues. The NHGRI-EBI GWAS catalog[16] was interrogated with the aim

of identifying possible pleiotropy with other cardiovascular phenotypes. At each locus, we defined a region based on LD span ($r^2 > 0.2$) with the lead SNP. We searched the GWAS catalog for all SNPs within these regions and report LD of proxies with the lead SNP when available. LD information was identified using the SNiPA tool[17] (Available at http://www.snipa.org. Accessed 6-24-2016.)

Expression Quantitative Trait Locus analyses

1. eQTL analyses in the Cleveland Clinic Atrial Tissue Bank and Arrhythmia Biorepository: We performed analyses of gene expression in human left atrial tissue samples obtained from the Cleveland Clinic Atrial Tissue Bank and Arrhythmia Biorepository. Genotypes were determined using the Illumina Human Hap550 v3 or Hap610 v1 chips; whereas RNA expression levels were determined using the Illumina HumanHT-12 v3 or v4 chips. The atrial samples were obtained from 289 individuals of European American (EA) ethnicity and 40 individuals of African American (AA) ethnicity. Of the EA individuals, 80 were female, 70 had no history of atrial fibrillation, and 136 were in atrial fibrillation at the time of tissue acquisition; 266 samples were from left atrial appendage (LAA) tissue and 23 the left atrial pulmonary vein junction tissue (LA-PV). Of the AA individuals, 25 were female, 16 had no history of atrial fibrillation, and 12 were in atrial fibrillation at the time of tissue acquisition; 34 samples were from LAA and 6 from LA-PV tissue. Methods have previously been described in depth by Deshmukh et al.[18] We performed cis-eQTL analyses for all statistically significant genetic variants identified in GWAS analyses. The Benjamini and Hochberg adjustment was applied to the results to control the false discovery rate (FDR).[19] P-values were adjusted based on the FDR of both genome-wide testing and specific variant sets, respectively. Probe-variant pairs with a genome-wide adjusted P-value less than 0.05 were deemed significant.

2. Examination of eQTLs in cardiac and skeletal muscle tissues from the GTEx database: The GTEx database was interrogated for all genetic loci associated with atrial fibrillation in the present meta-analyses. We selected the index variants and all proxies at the atrial fibrillation loci and looked for eQTLs in a subset of the GTEx database for right atrial, left ventricular, and skeletal muscle tissues that are most relevant to atrial fibrillation.

3. GTEx region based analyses were performed by comparing the percent of atrial fibrillation loci with at least one eQTL to the percent of control loci with at least one eQTL. All tissues in the GTEx database were used for this analysis. Atrial fibrillation loci and control loci were defined as described in the "Functional element enrichment" section above. Statistical significance was calculated by a one-tailed test based on 1,000 bootstrap samples from each set of control loci.

Replication of genetic variants specific to African American ancestry GWAS meta-analysis

We sought to replicate variants specific to the African American ancestry GWAS meta-analysis in 447 atrial fibrillation cases and 442 referents of African American ancestry. Custom TaqMan® genotyping probes for rs115339321 and rs79433233 were obtained from Life Technologies. Genotyping was performed on 5 ng of DNA input using the TaqMan® genotyping master mix on a Bio-Rad CFX384 real time PCR instrument. Genotyping was performed in 447 atrial fibrillation cases and 442 referents obtained from four studies (BioVU, Duke Biobank, MGH, and Penn Biobank), with genotype calls being performed by end state fluorescence after 40 cycles. See **Supplementary Results** and **Supplementary Tables 25-26** for further details.

In silico replication in the BioBank Japan (BBJ) study

The variant with the lowest P-value at each independent novel atrial fibrillation locus was selected for in silico replication in the results from GWAS analysis in 8180 individuals with atrial fibrillation and 28,612 referents from the BioBank Japan study. The cases were selected from the Biobank Japan which contains DNA and serum samples collected throughout Japan and atrial fibrillation was defined as persistent or paroxysmal atrial fibrillation diagnosed by a physician. The referents were selected from the Tohoku Medical Megabank organization,[20] the Japan Public Health Centre-based Prospective study, and the Japan Multi-institutional Collaborative Cohort (J-MICC) Study. Samples were genotyped using the Illumina Human OmniExpress BeadChip Kit and Infinium OmniExpressExome BeadChip Kit. Only autosomal variants were included in the GWAS. Variants with call rate <99%, variants that deviated from Hardy-Weinberg equilibrium among control samples ($<1x10^{-6}$), and non-polymorphic variants were excluded.

In silico replication in the UK Biobank study

Replication was performed using 143,218 unrelated adults of primarily European ancestry (>80%), aged 40-69 years old between 2006 and 2010, from the UK Biobank interim dataset released in May 2015. We defined atrial fibrillation as reported during a baseline interview; presence of a procedure code for cardioversion, atrial flutter or fibrillation ablation, or atrioventricular node ablation; billing code for atrial fibrillation; or atrial fibrillation reported on a death record (specific codes used in the definition are available upon request). Of the 143,218 individuals in the replication dataset, we identified 3366 individuals with atrial fibrillation, according to the criteria above. Details of genotyping, imputation, and calculation of principal components of ancestry in the UK biobank interim dataset can be found on the UK biobank website (http://www.ukbiobank.ac.uk/). Briefly, samples were genotyped either by UK BiLEVE Axiom array

(UKBL) or UK Biobank Axiom array (UKBB). Both arrays include ~800,000 SNPs and more than 95% of common marker contents are similar. Imputation was phased by modified version of SHAPEIT2 and imputed by IMPUTE2, using a combined panel of UK10K haplotype and 1000G phase 3 as the reference panel. All significant variants detected in the discovery study passed quality control filters in the UK biobank data (imputation quality info ≥ 0.4, variant missing rate < 5%, individual missing rate < 10%, and variant genotype probability > 0.9 in > 90% of the individuals). Variants were then transformed to hard-called genotypes (probability threshold ≥ 0.9, minor allele frequency (MAF) ≥ 0.01, and missing rate per variant <5%). We used logistic regression to test the association between each hard-called variant and risk of atrial fibrillation using an additive genetic model, adjusting for baseline age, sex, array, and the first 15 principal components of ancestry. Quality control, transformation and analyses were performed by QCTOOL and Plink v1.90b. Since we performed an in silico replication of 31 variants, we set a conservative significance threshold of 1.6 x 10-3 (0.05 / 31).

Pathway analyses

Pathway analyses provide a potential route to investigate the collective effects of multiple genetic variants on biological systems (see **Supplementary Results** and **Supplementary Tables 27-29**). We utilized two different methods for pathway analysis:

1. DEPICT

We ran the analysis DEPICT,[21] which integrates multiple layers of evidence to identify causal genes at GWAS loci. From meta-analysis results, we first performed clumping to identify independent loci using plink.[22] We then performed analysis using DEPICT with the default settings.

2. Ingenuity Pathway Analysis (IPA)

Data were analyzed through the use of QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). For each of the tested genetic variants, we mapped it back to the reference human genome (NCBI Build 37, 2009) and examined its location relative to RefSeq genes (May 15, 2016). The gene score was defined as the most significant variants that were located within 110kb upstream and 40kb downstream of the gene's most extreme transcript boundaries. Of the 27,011 genes evaluated, 338 reached a score less than $5x10^{-6}$. These genes were then imported into IPA analysis. Fisher's exact test was used to justify the enrichment of each of the canonical pathways.

Assessment of pleiotropy with the ischemic stroke phenotype

In order to evaluate pleiotropy with the ischemic stroke phenotype, we selected the variant with the lowest P-value at each independent novel atrial fibrillation locus and performed a lookup in the results from 1000 Genomes imputed GWAS meta-analyses from the Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium (4348 stroke patients and 80,613 referents)[23] and the METASTROKE consortium (10,307 ischemic stroke cases and 19,326 referents) of the International Stroke Genetics Consortium (ISGC).[24]

Online methods references

1. Benjamin, E. J. *et al.* Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat. Genet.* **41**, 879–81 (2009).
2. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
3. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M. J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385 (2008).
4. Hoggart, C. J., Clark, T. G., De Iorio, M., Whittaker, J. C. & Balding, D. J. Genome-wide significance for dense SNP and resequencing data. *Genet. Epidemiol.* **32**, 179–185 (2008).
5. Dudbridge, F. & Gusnanto, A. Estimation of significance thresholds for genomewide association scans. *Genet. Epidemiol.* **32**, 227–234 (2008).
6. Sham, P. C. & Purcell, S. M. Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* **15**, 335–346 (2014).
7. Lumley, T., Brody, J., Dupuis, J. & Cupples, A. Meta-analysis of a rare-variant association test. (2012).
8. Lee, S., Wu, M. C. & Lin, X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics* **13**, 762–775 (2012).
9. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
10. Pers, T. H., Timshel, P. & Hirschhorn, J. N. SNPsnap: a Web-based tool for identification and annotation of matched SNPs. *Bioinformatics* **31**, 418–20 (2015).

11. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).

12. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–5 (2013).

13. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).

14. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40**, D930-4 (2012).

15. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–11 (2001).

16. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42**, D1001-6 (2014).

17. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNiPA: An interactive, genetic variant-centered annotation browser. *Bioinformatics* **31**, 1334–1336 (2015).

18. Deshmukh, A. *et al.* Left atrial transcriptional changes associated with atrial fibrillation susceptibility and persistence. *Circ. Arrhythm. Electrophysiol.* **8**, 32–41 (2015).

19. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 289–300 (1995).

20. Kuriyama, S. *et al.* The Tohoku Medical Megabank Project: Design and Mission. *J. Epidemiol.* **26**, 493–511 (2016).

21. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).

22. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–75 (2007).

23. Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, Stroke Genetics Network (SiGN) & International Stroke Genetics Consortium (ISGC). Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. *Lancet. Neurol.* **15**, 695–707 (2016).

24. Malik, R. *et al.* Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology* **86**, 1217–26 (2016).

SUPPLEMENTARY NOTE
Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation

INDEX

4. Supplementary Results

Ancestry-specific GWAS meta-analyses

GWAS meta-analyses of incident and prevalent atrial fibrillation in Europeans

Replication of genetic variants specific to African American ancestry GWAS meta-analysis

Pathway analysis

1. DEPICT
2. IPA

5. Acknowledgments

6. Supplementary References

Supplementary Note - **Novel genetic loci for atrial fibrillation**

1. SUPPLEMENTARY TABLES

Supplementary Table 1. Baseline characteristics GWAS meta-analysis

Enclosed electronic excel file

Supplementary Table 2. Baseline characteristics for the exome chip meta-analysis

Enclosed electronic excel file

Supplementary Table 3. Detailed description of the genes at novel atrial fibrillation loci

Chromosomal location, Sentinel variant, *Gene(s)*: Description of the genes at the locus.
GWAS loci
1q24, rs72700118, *METTL11B/KIFAP3*: The most significant variant at 1q24 lies downstream of the closest gene, *METTL11B*, which encodes an N-terminal mono-methyltransferase that regulates DNA-protein interactions.[3] It is an important cell cycle regulator and mediator of DNA repair mechanisms since *METTL11B* knockout mice either die shortly after birth or display various developmental defects.[4] Interestingly, it also has been shown that METTL11B might act as a tumor suppressor protein in breast cancer.[5] *METTL11B* is highly expressed in right atrial and left ventricular tissue in GTEx. Analyses revealed that METTL11B may potentially interact with the atrial specific myosin light chain gene (MYL4) that has been linked to atrial fibrillation.[6,7] The locus also includes the gene *KIFAP3,* for which there also was a significant eQTL in the CCAF human atrial samples (**Supplementary Table S17 and S19**). *KIFAP3* encodes the kinesin associated protein 3, which regulates small G proteins by stimulating GDP/GTP exchange reactions or inhibiting their membrane interactions.[8] The gene is expressed in right atrial and left ventricular human tissue samples in the GTEx database. It is thought that this protein serves as a linker between human chromosome-associated polypeptide (HCAP) and KIF3A/B, a kinesin superfamily protein in the nucleus, and that this motor complex mediates binding to motor proteins enabling mainly anterograde transport of vesicles along microtubules.[9,10] *KIFAP3* variants have previously been associated with increased survival in sporadic amyotrophic lateral sclerosis and a combined phenotype of obesity and endometriosis in GWAS.[11,12] Reduced expression of *KIFAP3* has been demonstrated in clear cell renal carcinomas and was correlated with tumor aggressiveness and poorer patient outcomes,[13] whereas overexpression of the gene has been shown in breast cancer tumors.[14] In addition, *KIFAP3* has been shown to be involved in control of female puberty onset.[15] No relation to cardiac phenotypes have been noted for *KIFAP3* so far.
2p13, rs3771537, ***ANXA4/GMCL1:*** At 2p13, the most significant variant was intronic to *ANXA4*, whereas there were significant eQTLs for *ANXA4, GMCL1, PCYOX1,* and *SNRNP27* in GTEx left ventricle and skeletal muscle tissue (**Supplementary Table S17-S18**). *ANXA4* encodes Annexin 4, which is a Ca^{2+} and phospholipid binding protein that modulates membrane permeability, growth, apoptosis.[16] It has been demonstrated to be overexpressed in various cancers like lung cancer, colorectal cancer or prostate cancer where it enhances tumor invasion and chemotherapy resistance.[17] It has further been shown that *ANXA4* is involved in β-adrenergic signaling since Anxa4$^{-/-}$ mice show increased cellular cAMP levels and enhanced left ventricle contraction force upon adrenergic stimulation, whereas calcium stimulation in the left atrium lead to increased contraction force relative to wildtype mice.[18] Moreover, annexin 4 has been shown to bind to adenylyl cyclase type 5; thus, it has been suggested that annexin 4 directly modulates the β-adrenoceptor cAMP-dependent signal transduction pathway by inhibiting adenylyl cyclase 5.[18] In line with this hypothesis, *ANXA4* has been shown to be upregulated in human failing hearts.[19] *GMCL1,* which encodes Germ Cell-Less protein 1, is predominantly expressed in the testis, where it is

involved in spermatogenesis.[20,21] It has been demonstrated to regulate chromatin in germ cells by interacting with GAGE12I[22] and might also have a role in oncogenesis since it is expressed in various cancers like B cell lymphoma.[23] A direct link to cardiac physiology or disease; however, is currently missing.

2p14, rs2540949, *CEP68*: The most significant variant at 2p14 was intronic to *CEP68*, which encodes the centrosomal protein 68 that is important for the cell cycle by regulating centrosome cohesion.[24] There were significant eQTLs for *CEP68* in both the CCAF human atrial samples (**Supplementary Table S17 and 19**) and GTEx atrial, left ventricle, and skeletal muscle tissue (**Supplementary Table S17-S18**). At the onset of mitosis CEP68 dissociates from the centrosomes allowing the centrosomes to separate.[25] Variants in *CEP68* has been associated with aspirin-induced asthma[26] and acute urticaria/angioedema induced by non-steroidal anti-inflammatory drugs.[27]

2p31, rs2288327, *TTN/TTN-AS1:* At 2q31 we identified six significant coding variants in the A-band and M-line of titin, which all were predicted to be benign by PolyPhen and SIFT. The *TTN* gene spans 363 exons and the encoded protein stretches through half the length of a sarcomere.[28] Titin ensures sarcomere integrity and elasticity, and binds actin and myosin, which are crucial players in the contractile machinery in striated muscle.[29,30] Truncating mutations in titin have been shown to be the most important cause of dilated cardiomyopathy;[31–35] however, the gene displays considerable variation, making interpretation of mutational findings challenging.[36] Titin has been associated with the QT-interval in previous GWAS,[37,38] but the lead variant in our study (rs2288327) was not in LD with the QT-associated *TTN* variant (rs7561149, r^2=0.004).

5q22, rs337711, *KCNN2:* The variant at 5q22 is located in an intron of the gene *KCNN2* that encodes the small-conductance calcium-activated potassium channel, subfamily N, member 2 or SK2 channel. There was a significant eQTL for *KCNN2* itself in the CCAF human atrial tissue samples (**Supplementary Table S19**). This ion channel is predominantly expressed in the atria[39] and is involved in electrical remodeling resulting in atrial fibrillation.[39,40] In chronic atrial fibrillation, SK2 expression is reduced leading to significant changes in action potential duration (APD), a finding that has been confirmed in knockout mice. Furthermore, SK2 channels have been demonstrated to be involved in ventricular repolarization and also development of ventricular arrhythmias, especially in failing hearts where SK2 channels are upregulated both in patients and animal models.[41–45] Functional analysis revealed that the activation and modulation of SK2 channels is dependent on Ryr2-mediated calcium release[46] and that amiodarone can inhibit SK2 channels in a time- and voltage-independent but calcium-dependent mechanism, partly explaining its antiarrhythmic effects in failing hearts.[47] Additionally, genome-wide association studies have identified *KCNN2* as a susceptibility gene for coronary aneurysms in Kawasaki disease.[48,49] SK2 channels have also been shown to be involved in ischemia-induced neuronal cell death,[50,51] neuronal plasticity and learning,[52–55] drug addiction,[56,57] regulation of sleep duration,[58] and maintenance of the ionic milieu of the inner ear fluid.[59] They may be therapeutic targets for Parkinson's disease, since activation of SK2 channels provides protective effects in human dopaminergic neurons.[60]

5q31, rs2967791, *PKD2L2/KLHL3/WNT8A/FAM13B*: *PKD2L2* encodes the polycystic kidney disease 2-like 2 protein that belongs to the transient receptor potential (TRP) superfamily and is highly expressed in human brain, kidney, and testis.[61,62] In rodents, it is also expressed in the heart and has been demonstrated to be involved in calcium homeostasis, proliferation, and apoptosis.[61–63]

KLHL3 encodes the gene Kelch Like Family Member 3 that is part of the E3 ubiquitin ligase complex regulating the sodium/chloride cotransporter (NCC), the epithelial sodium channel (ENaC), and the renal outer medullary potassium channel (ROMC) in the kidney.[64,65] It is an important regulator of the electrolyte homeostasis and therefore the blood pressure.[66,67] Genetic variants of *KLHL3* have been described to cause familial hyperkaliemic hypertension. [65,68,69]

WNT8A is a member of the WNT/beta catenin-signaling network that plays an essential role in development and carcinogenesis.[70] *WNT8A* has been demonstrated to regulate body axis extension[71] and neuroectodermal posteriorization.[72] *WNT8A* polymorphisms have been shown to be associated with Hirschsprung's disease and its expression is upregulated in stenotic colon segments in patients.[73] Interestingly, *in vitro* overexpression of WNT8 results in impaired calcium handling[74] and might therefore also be involved in atrial fibrillation pathophysiology.

For the 5q31 locus, we identified an eQTL for the gene *FAM13B* in eQTL enrichment analysis (**Supplemental table S17**). *FAM13B* (syn. *C5ORF5*) consists of 23 exons spanning over 27 kb; the transcript is 5.47 kb and encodes a protein of 915 amino acids.[75] It contains a putative rhoGAP domain at the N-terminus and two bipartite nuclear localization signals and is predominantly expressed in brain and male reproductive tissue[76] (Human Protein Atlas available from www.proteinatlas.org). So far, *FAM13B* has not been reported in a cardiovascular context.

8p22, rs7508, *ASAH1/PCM1*: At 8p22, the lead atrial fibrillation risk variant was associated with decreased expression of *ASAH1* (rs7508; $P = 5.1 \times 10^{-3}$) in CCAF human atrial samples and increased expression of *PCM1* (rs7508; $P = 9.6 \times 10^{-14}$) in both the CCAF samples (**Supplementary Table S17 and S19**) and GTEx left ventricle and skeletal muscle tissue (**Supplementary Table S17-S18**). *ASAH1* encodes the acid ceramidase 1 that is involved in lipid metabolism by degradation of ceramide into sphingosine and free fatty acids within lysosomes.[77,78] Overexpression of ceramidase has been reported in several cancer cell types,[79–81] resulting in increased proliferation[82] and invasiveness,[83,84] predominantly in prostate cancer, which in turn has led to studies showing promising results of ceramidase inhibitors as new cancer therapeutics.[85,86] Ceramidase has also been implicated in Farber's disease (lipogranulomatosis),[87,88] spinal muscular atrophy with myoclonic epilepsy,[89] and Alzheimer's disease.[90] *ASAH1* is highly expressed in the heart.[91] Accumulation of ceramide has been shown to result in oxidative stress, electron transport chain dysfunction, and cardiomyocyte apoptosis in rats.[92,93]

PCM1 encoding pericentriolar material 1, has been demonstrated to be an integral component of centriolar satellites in ciliogenesis.[94] It has also been shown to be involved in neurogenesis ,[95] the centrosomal actin network,[96] hematological neoplasms,[97] and associated with schizophrenia.[98]

10q24, rs35176054, *SH3PXD2A*: The variant at 10q24 is located intronic to the gene *SH3PXD2A* that encodes the SH3 and PX domain-containing protein 2A or Adapter protein TKS5 that plays an essential role in various malignancies. It interacts with Src tyrosine kinase to promote tumor growth and the formation of invadopodia resulting in degradation of extracellular matrix and invasion of cancer cells into surrounding tissue in breast, ovarian, colon, lung, prostate cancer, melanoma, and glioma.[99–103] Its expression level has been demonstrated to be negatively correlated with tumor size and patient survival in ovarian cancer.[104,105] However, it is also involved in normal embryonic development by regulating neural crest migration[106,107] and in macrophage or microglia physiology.[108,109]

11q24, rs75190942, *KCNJ5:* The genetic variant rs75190942 is located at 11q24 within the gene *KCNJ5*, that encodes the G protein-activated inward rectifier potassium channel 4 (Kir3.4/GIRK4). There was a significant eQTL for *KCNJ5* itself in CCAF human atrial tissue samples (**Supplementary Table S17 and S19**) and in GTEx left ventricle tissue (**Supplementary Table S17-S18**). GIRK4 is known to form heteromeres with Kir3.1/GIRK1/KCNJ3, constituting the I_{KAch} channel complex, which contributes to the regulation of the membrane potential in the sinoatrial node and atria – making it a therapeutic target for atrial fibrillation. This ion channel has been shown to regulate pacemaker activity and recovery of resting heart rate after sympathetic stimulation.[110] GIRK4 inactivation can also rescue arrhythmias that are induced by genetic silencing of funny currents.[111] Furthermore, it determines inducibility, dynamics and termination of atrial fibrillation by regulating action potential duration.[112] Additionally, genetic polymorphisms in *KCNJ5* are associated with early-onset lone atrial fibrillation,[113] whereas mutations in this gene have been shown to cause long QT syndrome.[114] GIRK4 is also expressed in the ventricles and contributes to ventricular repolarization[115] and has been shown to be significantly downregulated in patients with dilated cardiomyopathy.[116] Furthermore, mutations in *KCNJ5* can cause Andersen-Tawil syndrome,[117] primary aldosteronism[118] and has been detected in adrenal tumors.[119] Also, *KCNJ5* is associated with Tourette Syndrome and Attention-Deficit/Hyperactivity Disorder.[120]

ExWAS loci

3p22, rs6800541, *SCN10A:* The variant rs6800541 is located intronic to *SCN10A*, the gene that encodes the sodium channel Nav1.8. It is highly expressed in primary sensory neurons and dorsal root ganglion neurons and has been linked to nociception, painful neuropathy, and multiple sclerosis.[121] Recently, it has been shown that Nav1.8 is also expressed in the heart where it contributes to the late sodium current.[122,123] Genome-wide association studies demonstrated genetic variants in *SCN10A* as risk loci for quantitative ECG traits like PR interval,[124–128] and QRS duration,[126,129,130] as well as for atrial fibrillation[124,126,130,131] and Brugada Syndrome.[132] Also, mutations in *SCN10A* has been shown to be responsible for a large fraction of cases of Brugada Syndrome.[133] Data suggest that *SCN10A* affects cardiac conduction either directly through cardiomyocytes, indirectly through intracardiac neurons, or by modulation of *SCN5A* expression.[134,135]

12p12, rs11047543, *SOX5:* The most significant SNP at 12p12 is located downstream of the *SOX5* gene. SOX5 is a transcription factor that has been shown to be involved in limb development,[136] chondrogenesis,[137] brain development,[138] and lung development.[139] Our current study confirmed previous genome-wide association studies that showed a significant association between *SOX5* and early-onset atrial fibrillation.[124,140] Furthermore, *SOX5* has been demonstrated to be significantly associated with PR-interval,[124] left ventricular mass,[141] resting heart rate,[142] osteoporosis,[143] systemic sclerosis,[144] AIDS,[145] chronic obstructive pulmonary disease,[139] and non-obstructive azoospermia.[146] Additionally, it is involved in the development of lung cancer,[147] hepatocellular carcinoma,[148] follicular lymphoma,[149] and melanoma.[163]

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Locus identified in both GWAS and EWAS:
6q22, rs4946333 (GWAS), rs89107 (EWAS), *SLC35F1/PLN:*** At 6q22 we identified a locus including the phospholamban gene (*PLN*), *SLC35F1,* and *CEP85L*. Phospholamban regulates cardiac contractility and relaxation through inhibiting the cardiac muscle sarcoplasmic reticulum calcium ATPase SERCA.[164] Mutations in this gene has been associated with hypertrophic[165,166] and dilated cardiomyopathy.[167,168] *SLC35F1* encodes a member of the solute carrier family 35. *SLC35F1* knockout mice display reduced levels of hemoglobin and lactate dehydrogenase but do not show any further phenotype. Previous GWAS have associated the locus surrounding *SLC35F1/PLN/CEP85L* with resting heart rate,[6,15] QT-interval,[12,14] and left ventricle internal diastolic diameter.[11] One of the variants associated with heart rate by den Hoed et al. also associated with atrial fibrillation in secondary analyses.[6]

Supplementary Table 4. Results from Asian ancestry SKAT gene based test

Gene	Chr	CMAF	N variants	P-value
Filter: Variants predicted to be damaging				
SH3PXD2A	10q24	0.4	6	4.77×10^{-11}
Filter: Nonsynonymous and splice site variants				
SH3PXD2A	10q24	0.4	11	4.21×10^{-11}

Chr, chromosome; CMAF, cumulative minor allele frequency per gene.

Supplementary Table 5. Single variant association results for the variants that were analyzed in the two significant gene-based tests for *SH3PDX2A* in the Asian ancestry group.

rsID	Risk/ref allele	Amino acid substitution**	RAF, %	OR	95% CI	P-value
rs149867987	A/G	p.His110Tyr	0.01	16.72	2.23-125.31	0.006
rs200938753*	G/A	p.Arg761Cys	99.89	1.45	0.74-2.84	0.27
rs202011870*	C/A	p.Leu396Arg	0.18	4.68	2.97-7.39	3.30E-11
rs201065560*	A/G	p.Arg1031Cys	0.02	2.03	0.55-7.47	0.29
rs74661743*	G/A	p.Arg1003Cys	99.93	1.02	0.42-2.47	0.97
rs79061932	G/A	p.Arg994Cys	99.99	1.13	0.07-18.44	0.93
rs201439736	C/T	p.Ala886Thr	99.97	1.44	0.46-4.52	0.54
rs201054626*	T/C	p.Arg302Gln	0.01	4.85	0.83-28.47	0.08
rs143819462	T/C	p.Arg269Gln	0.01	2.34	0.39-13.93	0.35
rs147297499	T/C	p.Asp231Asn	0.005	13.31	0.67-264.24	0.09
rs143409187*	T/C	p.Arg102Gln	0.007	2.85	0.15-55.03	0.49

The gene-based test was significant for the subset of nonsynonymous and splice site variants, which included all listed variants, and the subset of nonsynonymous possibly damaging variants, which included 6 of the listed variants (*). **NCBI Reference sequence accession and version number NP_055446.2. RAF, risk allele frequency; CI, confidence interval; OR, odds ratio.

Supplementary Table 6. Results from ancestry-specific GWAS meta-analyses

	rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
EUR	*15,993 cases, 113,719 referents*								
	Novel associations								
	rs10800507	1q24	*METTL11B/KIFAP3*	Intergenic	C/G	51	1.09	1.06-1.12	$1.87x10^{-11}$
	rs62133983	2p13	***ANXA4****/GMCL1*	Intronic	G/T	52	1.09	1.06-1.12	1.36x10-10
	rs2723064	2p14	*CEP68*	Intergenic	T/C	61	1.09	1.06-1.12	1.88x10-10
	rs6864727	5q31	***PKD2L2****/WNT8A/FAM13B*	Intronic	C/T	31	1.08	1.05-1.11	$1.12x10^{-8}$
	rs281868	6q22	***SLC35F1****/PLN*	Intronic	G/A	50	1.08	1.05-1.10	$1.03x10^{-8}$
	rs7508	8p22	***ASAH1****/PCM1*	3'UTR	A/G	73	1.10	1.06-1.13	$6.34x10^{-10}$
	rs35176054	10q24	***SH3PXD2A***	Intronic	A/T	13	1.14	1.10-1.18	$1.75x10^{-11}$
	rs75190942	11q24	***KCNJ5***	Intronic	A/C	8	1.18	1.11-1.25	$2.82x10^{-8}$
	rs2921421	15q21	*CGNL1*	Intergenic	G/C	3	1.72	1.42-2.09	$3.29x10^{-8}$
	Previously known associations								
	rs11264280	1q21	*KCNN3*	Intergenic	T/C	32	1.13	1.10-1.16	$2.77x10^{-17}$
	rs651386	1q24	*PRRX1*	Intergenic	A/T	57	1.11	1.08-1.14	$6.23x10^{-15}$
	rs2129977	4q25	*PITX2*	Intergenic	A/G	22	1.45	1.41-1.49	$7.25x10^{-136}$
	rs12664873	6q22	*GJA1*	Intergenic	T/G	69	1.08	1.05-1.12	$1.80x10^{-8}$
	rs11773845	7q31	***CAV1/2***	Intronic	A/C	60	1.10	1.07-1.13	$3.35x10^{-13}$
	rs7026071	9q22	***C9orf3***	Intronic	T/C	41	1.09	1.07-1.12	$2.86x10^{-11}$
	rs10824026	10q22	*SYNPO2L*	Intergenic	A/G	84	1.13	1.09-1.17	$8.29x10^{-11}$
	rs11598047	10q24	***NEURL1***	Intronic	G/A	17	1.18	1.14-1.22	$3.16x10^{-21}$
	rs883079	12q24	***TBX5***	3'UTR	T/C	72	1.11	1.08-1.15	$1.31x10^{-13}$
	rs7183206	15q24	*HCN4*	Intergenic	A/G	15	1.13	1.09-1.18	$7.70x10^{-12}$
	rs2106261	16q22	***ZFHX3***	Intronic	T/C	18	1.19	1.15-1.23	$4.01x10^{-24}$
AA	*641 cases, 4956 referents*								
	rs6843082	4q25	*PITX2*	Intergenic	G/A	30	1.40	1.24-1.58	$4.31x10^{-8}$
AS	*837 cases, 2456 referents*								
	Novel association								
	rs7138621	12q15	*CPSF6*	Intergenic	G/C	95	7.92	4.26-14.73	$6.48x10^{-11}$
	Previously known association								
	rs2723334	4q25	*PITX2*	Intergenic	T/C	70	1.94	1.68-2.25	$8.46x10^{-19}$

The most significant variant at each genetic locus associated with atrial fibrillation is listed. Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed. Chr, chromosome; CI, confidence interval; OR, odds

ratio; EUR, European ancestry; AA, African American ancestry; AS, Asian ancestry; RAF, risk allele frequency.

Supplementary Table 7. Results from European and Asian ancestry ExWAS meta-analysis

	rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
EUR	*13,496 cases, 96,273 referents*								
	Novel associations								
	rs6800541	3p22	***SCN10A***	Intronic	T/C	60	1.08	1.05-1.12	8.75×10^{-7}
	rs89107	6q22	***SLC35F1****/PLN*	Intronic	G/A	50	1.09	1.06-1.13	2.71×10^{-7}
	rs11047543	12p12	*SOX5*	Intergenic	G/A	85	1.13	1.08-1.18	4.65×10^{-7}
	Previously known associations								
	rs13376333	1q21	***KCNN3***	Intronic	T/C	31	1.14	1.10-1.17	1.58×10^{-13}
	rs2200733	4q25	*PITX2*	Intergenic	T/C	12	1.60	1.52-1.67	9.95×10^{-90}
	rs3807989	7q31	***CAV1***	Intronic	G/A	59	1.09	1.06-1.13	2.93×10^{-8}
	rs60632610	10q22	***SYNPO2L***	Exonic; nonsyn	C/T	85	1.13	1.08-1.18	2.53×10^{-8}
	rs2106261	16q22	***ZFHX3***	Intronic	A/G	17	1.21	1.16-1.26	3.37×10^{-18}
AS	*8180 cases, 28,612 referents*								
	Novel associations								
	rs55952639	2p14	***CEP68***	Exonic; syn	T/C	76	1.13	1.07-1.18	1.29×10^{-6}
	rs11047543	12p12	*SOX5*	Intergenic	G/A	88	1.18	1.10-1.26	1.16×10^{-6}
	Previously known associations								
	rs17042171	4q25	*PITX2*	Intergenic	A/C	48	1.69	1.62-1.76	4.04×10^{-137}

The most significant variant at each genetic locus associated with atrial fibrillation is listed. Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants the closest gene(s) are listed. Chr, chromosome; CI, confidence interval; OR, odds ratio; EUR, European ancestry; AA, African American ancestry; AS, Asian ancestry; nonsyn, nonsynonymous; syn, synonymous; RAF, risk allele frequency

Supplementary Table 8. Results from European incident atrial fibrillation GWAS meta-analysis

rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
rs11264280	1q21	*KCNN3*	Intergenic	T/C	32	1.12	1.08-1.16	$3.57x10^{-9}$
rs6843082	4q25	*PITX2*	Intergenic	G/A	21	1.38	1.33-1.44	$8.21x10^{-57}$
rs7394190	10q22	*SYNPO2L*	Intergenic	G/A	84	1.15	1.09-1.21	$3.09x10^{-8}$
rs60848348	10q24	***NEURL1***	Intronic	T/C	20	1.13	1.09-1.18	$1.69x10^{-8}$
rs4499262	16q22	***ZFHX3***	Intronic	A/C	17	1.14	1.09-1.19	$4.01x10^{-8}$

The most significant variant at each genetic locus associated with atrial fibrillation is listed. Gene names in bold font indicate that the variant is located within the gene. Chr, chromosome; CI, confidence interval; OR, odds ratio; RAF, risk allele frequency

Supplementary Table 9. Results from European prevalent atrial fibrillation GWAS meta-analysis

rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
Novel associations								
rs72700118	1q24	*METTL11B/KIFAP3*	Intergenic	A/C	11	1.24	1.17-1.31	$9.93x10^{-13}$
rs6546550	2p13	***ANXA4****/GMCL1*	Intronic	C/G	54	1.12	1.08-1.16	$1.36x10^{-8}$
rs1454934	12p11	***PKP2***	Intronic	T/C	16	1.16	1.1-1.22	$4.18x10^{-8}$
Previously known associations								
rs36004974	1q21	***KCNN3***	Intronic	G/A	32	1.14	1.1-1.19	$4.36x10^{-10}$
rs577676	1q24	*PRRX1*	Intergenic	C/T	55	1.15	1.1-1.19	$2.77x10^{-12}$
rs61303432	4q25	*PITX2*	Intergenic	T/C	14	1.71	1.62-1.8	$6.66x10^{-92}$
rs2109514	7q31	*CAV1/2*	Intergenic	A/G	51	1.15	1.11-1.19	$6.73x10^{-13}$
rs11598047	10q24	***NEURL1***	Intronic	G/A	17	1.24	1.18-1.31	$4.34x10^{-16}$
rs2106261	16q22	***ZFHX3***	Intronic	T/C	18	1.25	1.19-1.31	$9.68x10^{-20}$

The most significant variant at each genetic locus associated with atrial fibrillation is listed. Gene names in bold font indicate that the variant is located within the gene. Chr, chromosome; CI, confidence interval; OR, odds ratio; RAF, risk allele frequency

Supplementary Table 10. Comparison of results for common variant loci between the AFGen Consortium combined ancestry analysis and the Biobank Japan study.

Enclosed electronic excel file

Supplementary Table 11. Replication of the common variant loci identified in the AFGen Consortium combined ancestry analysis and the UK Biobank study.

Enclosed electronic excel file

Supplementary Table 12. Approximate and joint conditional analysis in European ancestry GWAS meta-analysis identify 20 independent genetic loci associated with atrial fibrillation

rsID	Chr	Gene	Location relative to gene	P-value
rs11264280	1	*KCNN3*	Intergenic	2.77×10^{-17}
rs10800507	1	*METTL11B*	Intergenic	1.87×10^{-11}
rs651386	1	*PRRX1*	Intergenic	6.23×10^{-15}
rs2723065	2	*CEP68*	Intergenic	1.91×10^{-10}
rs62133983	2	***ANXA4***	Intronic	1.36×10^{-10}
rs2129977*	4	*PITX2*	Intergenic	7.25×10^{-136}
rs6864727	5	***PKD2L2***	Intronic	1.12×10^{-8}
rs281868	6	***SLC35F1***	Intronic	1.03×10^{-8}
rs7773091	6	*GJA1*	Intergenic	2.02×10^{-8}
rs11773845	7	***CAV1***	Intronic	3.35×10^{-13}
rs7508	8	***ASAH1***	3'UTR	6.34×10^{-10}
rs7026071	9	***C9orf3***	Intronic	2.86×10^{-11}
rs11598047	10	***NEURL1***	Intronic	3.16×10^{-21}
rs35176054	10	***SH3PXD2A***	Intronic	1.75×10^{-11}
rs10824026	10	*SYNPO2L*	Intergenic	8.29×10^{-11}
rs75190942	11	***KCNJ5***	Intronic	2.82×10^{-8}
rs883079	12	***TBX5***	3'UTR	1.31×10^{-13}
rs2921421	15	*CGNL1*	Intergenic	3.29×10^{-8}
rs8040533	15	*HCN4*	Intergenic	3.09×10^{-11}
rs2106261	16	***ZFHX3***	Intronic	4.01×10^{-24}

Chr, chromosome; UTR, untranslated region. Bold font indicates that the variant lies within the gene. *The 4q25/*PITX2* region was not analyzed because the complexity of this association signal is not accurately evaluated with the GCTA method (**Online Methods**).

Supplementary Table 13. Overlap with atrial fibrillation risk factor GWAS loci

rsID	Chr	Closest gene*	rsID GWAS Catalog	LD	GWAS P-Value	HR	PR-S	PR-I	QRS	QT	Echo LVIDD	Stroke
ALL ANCESTRIES												
rs6843082	4	*PITX2(dist=154788);C4orf32(dist=1348486)*	rs6843082	1	$3.41x10^{-155}$							3
rs6843082	4	*PITX2(dist=154788);C4orf32(dist=1348486)*	rs12646447	0.51	$1.12x10^{-148}$							4
rs6843082	4	*PITX2(dist=154788);C4orf32(dist=1348486)*	rs2200733	0.51	$2.32x10^{-150}$							5
rs2967791	5	*KLHL3/WNT8A*	rs7722600	0.15	$1.25x10^{-6}$	6						
rs4946333	6	*SLC35F1/PLN*	rs457162	<0.10	0.0686					7		
rs4946333	6	*SLC35F1/PLN*	rs11752626	0.43	0.0001					8		
rs4946333	6	*SLC35F1/PLN*	rs11970286	0.48	$3.29x10^{-5}$					9,10		
rs4946333	6	*SLC35F1/PLN*	rs12210810	<0.10	0.001					10		
rs4946333	6	*SLC35F1/PLN*	rs12210733	<0.10	0.001					7		
rs4946333	6	*SLC35F1/PLN*	rs89107	0.99	$4.03x10^{-9}$						11	
rs4946333	6	*SLC35F1/PLN*	rs3902035	<0.10	0.002					7		
rs4946333	6	*SLC35F1/PLN*	rs11756438	0.29	0.0008					12		
rs4946333	6	*SLC35F1/PLN*	rs6906287	0.38	$5.84x10^{-5}$				13			
rs4946333	6	*SLC35F1/PLN*	rs11153730	0.45	$2.01x10^{-5}$	6				7,14		
rs4946333	6	*SLC35F1/PLN*	rs281868	1	$2.12x10^{-9}$	15						
rs1997572	7	*CAV1*	rs3807989	0.93	$1.47x10^{-14}$		16	9,17,18	9			
rs1997572	7	*CAV1*	rs11773845	0.94	$7.53x10^{-15}$			19,20				
rs1997572	7	*CAV1*	rs9920	0.15	0.0005					7		
rs883079	12	*TBX5*	rs883079	1	$1.80x10^{-15}$				21			
rs883079	12	*TBX5*	rs7312625	0.90	$1.03x10^{-14}$			22				
rs883079	12	*TBX5*	rs1895585	0.83	$1.25x10^{-14}$			19				
rs883079	12	*TBX5*	rs7135659	0.88	$9.59x10^{-15}$			20				
rs883079	12	*TBX5*	rs3825214	0.59	$1.82x10^{-10}$			9	9	9		
rs74022964	15	*HCN4(dist=15659);C15orf60(dist=58235)*	rs4489968	0.77	$4.59x10^{-11}$	6						
rs2106261	16	*ZFHX3*	rs879324	0.91	$1.13x10^{-25}$							3

Table showing overlap of genetic associations between cardiac phenotypes, identified through interrogation of the NHGRI-EBI GWAS catalog.[2] Numbers in superscript in the phenotype columns indicate references to the literature. Chr, chromosome; LD, linkage disequilibrium r^2 with lead SNP; HR, heart rate; PR-S, PR-segment; PR-I, PR-interval; LVIDD, Left Ventricle Internal Diastolic Diameter. *For intronic variants, the gene the variant is located within is listed; for intergenic variants, the closest genes upstream and downstream are listed.

Supplementary Table 14. Association between novel atrial fibrillation loci and stroke subtypes in the Neuro-CHARGE Stroke Consortium

			All stroke		Ischemic stroke		Cardioembolic stroke	
rsID	**Gene***	**Risk/ref allele**	**OR**	**P-value**	**OR**	**P-value**	**OR**	**P-value**
rs72700118	*METTL11B*	A/C	1.01	0.70	1.02	0.61	1.09	0.38
rs3771537	***ANXA4****/GMCL1*	A/C	1.00	0.85	0.99	0.75	0.99	0.88
rs2540949	***CEP68***	A/T	1.04	0.12	1.05	0.09	1.14	0.02
rs2288327	***TTN/TTN-AS1***	G/A	1.05	0.08	1.08	0.02	1.22	0.01
rs337711	***KCNN2***	T/C	0.97	0.16	0.96	0.18	0.97	0.63
rs2967791	***KLHL3****/WNT8A/FAM13B*	T/C	1.03	0.14	1.04	0.10	1.11	0.05
rs4946333	***SLC35F1****/PLN*	G/A	0.97	0.21	0.97	0.18	0.97	0.58
rs7508	***ASAH1***	A/G	1.04	0.12	1.04	0.17	1.11	0.14
rs35176054	***SH3PXD2A***	A/T	1.03	0.38	1.01	0.77	1.07	0.44
rs75190942	***KCNJ5***	A/C	1.01	0.85	1.04	0.45	-	-

OR, odds ratio. *Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed.

Supplementary Table 15. Association between novel atrial fibrillation loci and stroke subtypes in the Metastroke Consortium

			Ischemic stroke		Cardioembolic stroke		Large vessel disease		Small vessel disease	
rsID	**Gene***	**Risk/ref allele**	**OR**	**P-value**	**OR**	**P-value**	**OR**	**P-value**	**OR**	**P-value**
rs72700118	*METTL11B/ KIFAP3*	A/C	1.07	0.02	1.14	0.02	1.01	0.92	1.04	0.53
rs3771537	***ANXA4****/GMCL1*	A/C	0.99	0.52	1.02	0.57	0.94	0.08	1.00	0.95
rs2540949	***CEP68***	A/T	0.99	0.63	1.03	0.40	1.05	0.18	0.97	0.54
rs2288327	***TTN/TTN-AS1***	G/A	1.02	0.54	1.03	0.61	1.02	0.66	1.07	0.21
rs337711	***KCNN2***	T/C	1.01	0.50	1.08	0.04	1.00	0.90	0.94	0.19
rs2967791	***KLHL3****/WNT8A/FAM13B*	T/C	1.02	0.39	1.05	0.19	1.06	0.15	0.92	0.05
rs4946333	***SLC35F1****/PLN*	G/A	0.98	0.26	0.91	0.01	0.89	0.003	1.01	0.79
rs7508	***ASAH1***	A/G	0.98	0.37	1.00	1.00	1.03	0.45	0.94	0.17
rs35176054	***SH3PXD2A***	A/T	1.01	0.67	1.07	0.25	0.96	0.46	1.10	0.13
rs75190942	***KCNJ5***	A/C	1.02	0.59	1.09	0.31	1.03	0.73	0.98	0.80

OR, odds ratio. *Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed.

Supplementary Table 16. GO Terms Enriched in Atrial Fibrillation Associated Loci Compared to GWAS Catalog Loci and to 1000 Genomes Matched Loci

Gene Ontology Description	P-value	FDR Q-value
Compared to 1000 Genomes Matched Loci		
Small conductance calcium-activated potassium channel activity	9.48×10^{-5}	3.01×10^{-1}
Metal ion transport	1.62×10^{-4}	1.00
Potassium channel activity	2.52×10^{-4}	4.00×10^{-1}
Z disc	2.70×10^{-4}	3.85×10^{-1}
Monovalent inorganic cation transport	3.52×10^{-4}	1.00
Potassium ion transmembrane transport	5.08×10^{-4}	1.00
Cellular potassium ion transport	5.08×10^{-4}	1.00
Potassium ion transmembrane transporter activity	5.70×10^{-4}	6.04×10^{-1}
Regulation of cardiac muscle contraction	6.92×10^{-4}	1.00
Striated muscle tissue development	6.92×10^{-4}	1.00
Potassium ion transport	7.08×10^{-4}	1.00
Cation transport	7.34×10^{-4}	1.00
Regulation of heart rate	9.10×10^{-4}	1.00
Compared to GWAS catalog Loci		
Small conductance calcium-activated potassium channel activity	2.64×10^{-4}	7.43×10^{-1}
Z disc	2.67×10^{-4}	3.34×10^{-1}
Metal ion transport	3.17×10^{-4}	1.00
Potassium channel activity	4.14×10^{-4}	5.83×10^{-1}
Monovalent inorganic cation transport	7.01×10^{-4}	1.00

Supplementary Table 17. Summary of top eQTLS within atrial fibrillation associated loci

Enclosed electronic excel file

Supplementary Table 18. *In silico* eQTL analysis in GTEx database

Enclosed electronic excel file

Supplementary Table 19. eQTL analysis of in CCAF human atrial tissue samples

Enclosed electronic excel file

Supplementary Table 20. *In silico* functional evaluation of novel and replicated loci from GWAS and ExWAS combined ancestry analysis

Enclosed electronic excel file

Supplementary Table 21. Per study overlap of samples between GWAS and exome chip analyses

Study	Overlap	
	Cases	Controls
BioVU	206	3811
WGHS	934	20,266
FHS - incident	411	1612
FHS - prevalent	181	2123
CHS - incident	922	1979
CHS -prevalent	60	2900
AGES	354	2989
RS	346	2370
CAMP	665	2128
SHIP	99	2710
AFLMU/KORA	349	415
MGH	333	0
ARIC EA	1253	3415
ARIC AA	233	742
MESA	155	2372
GS:SFHS	203	6651
BioMe EA	290	857
BioMe AA	166	2041
BioMe HA	255	2800
BEAT-AF	1520	1516
BBJ	782	0
Total	9717	63,697

Supplementary Table 22. GWAS information per study

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
AFLMU/KORA	[169]	Illumina HumanCNV370 + Illumina Human550K	BeadStudio	≥98%	$<10^{-5}$	-	-	>1%	P<0.05	1	306,838	SHAPEIT v2.r790 + IMPUTE v.2.1.2	SNPTEST v2.5	7,540,650	1.023
AGES	[170]	Illumina HumanCNV370-Duo BeadChip	BeadStudio	≥97%	$<10^{-6}$	-	-	≥1%	P<0.05	0	329,804	MaCH v.1.0.16 + minimac	ProbABEL, R	I: 7,602,716 P: 6,085,662	I: 1.068 P: 1.006
ANGES	[171]	Illumina Metabochip	GenomeStudio	≥95%	$\geq 10^{-6}$	-	>3.18 SD from the mean removed	-	first 4 PCs	4	121,545	SHAPEIT v.2.r790 + IMPUTE2 v.2.3.0	SNPTEST v2.4.1	5,861,502	P&I: 1.011
ARIC	[172,173]	Affymetrix 6.0	Birdseed	≥95%	$<10^{-5}$	-	-	EA: >0.5% AA: >1%	Analysis committee recommendations	EA: 4 AA: 10	EA: 711,589 AA: 806,416	(1) Pre-phasing with ShapeIt (v1.r532) (2) Imputation with IMPUTE2.1.0	FAST	EA: 9,428,893 AA: 8,978,558	EA: 1.011 AA: 0.991
Beat-AF	[174]	Illumina HumanCoreExome	BeadStudio	≥95%	$>10^{-6}$	-	>3 SD from the mean removed	≥1%	First 10 PCs	10	254,488	SHAPEIT v2.r790 + IMPUTE v.2.3.2	SNPTEST v.2.5	9,309,201	1.022

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
BBJ	[175]	Illumina Human610 Quad and Illumina Human Hap550v3 BeadChip	Beadstudio	≥99%	$>10^{-6}$	-	-	≥1%	First 2 PCs	2	432,042	MaCH + minimac	PLINK v1.07	6,429,092	1.024
BioMe	[176]	Illumina HumanOmni ExpressExome-8 v1.0	zCall (GenomeStudio)	≥90%	$p>10^{-6}$	-	-	≥1%	first 4 PCs	4	768,517	IMPUTE2	SNPTEST v.2.5	EA: 7,022,478 AA: 8,200,353 HA: 8,139,248	EA: 1.008 AA: 1.019 HA: 1.026
BioVU	[177]	Illumina Omni5 + Omni1 + 1M + 660K	GenomeStudio	≥98%	$<10^{-5}$	-	-	≥1%	First 2 PCs	2	4,167,400	IMPUTE2 v2.3.0	PLINK v1.90	660: 3,187,278 omni: 4,373,169	660: 1.003 Omni: 1.01
CCAF	[169]	Hap550 v1&v3 chip + Hap610 v1 chip	BeadStudio	≥95%	FDR>10^{-4}	-	FDR>0.01	≥1%	P<0.05	4	516,461	Shapeit v2.r727 + IMPUTE v.2.3.0	SNPtest v.2.5	8,122,372	1.026
CHS - AA	[178]	HumanOmni1-Quad_v1	GenomeStudio	≥97%	$\geq 10^{-5}$	≤1 in CEPH trios	-	>0.01 %	PCs with P<0.05 and all PCs before the associated PC	3	963,248	IMPUTE version 2.2.2	R	8,152,032	1.001
CHS - EA	[178]	Illumina 370 CNV + ITMAT-Broad-CARe (IBC) Illumina iSELECT chip	BeadStudio	≥97%	$\geq 10^{-5}$	≤2 in CEPH trios	-	>0.01 %	PCs with P<0.05 and all PCs before the associated PC	0	359,592	MaCH + minimac	R	8,278,530	1.045

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
COROGENE	[179]	Illumina Metabochip + CoreExome	GenomeStudio	≥95%	$\geq 10^{-5}$	-	-	≥1%	-	0	553,581	IMPUTE v2.2.2	SNPTEST v2.4.1	6,956,681	1.019
FHS	[180,181]	Affymetrix, Gene Chip®, 500K Array Set & 50K Human Gene Focused Panel	BRLMM	≥97%	$<10^{-6}$		Subject heterozygosity >5 SD away from the mean	≥1%	All PCs u-ssociated, p>0.05	0	385,958	Mach1 v1.0.15	R packages kinship, GEE, COXPH	I: 7525764 P: 6556225	I: 1.019 P: 1.04
FINCAVAS	[182]	Illumina Metabochip + CoreExome	GenomeStudio	≥95%	$\geq 10^{-6}$	-	>3.23 SD from the mean removed	-	First 4 PCs	4	Metabochip: 120,689 CoreExome: 277,211	SHAPEIT v.2.r790 + IMPUTE2 v.2.3.0	SNPTEST v2.4.1	8,384,365	P&I: 1.04
GS:SFHS	[183]	Illumina Omni Express Plus Exome	BeadStudio	Omni ≥98% Exome ≥99%	$<10^{-6}$	-	-	Omni <1% Exome <0.01 %	PCs associated after adjustment for sex and age with p<0.05)	1	706,198 (690,759 Autosomes)	ShapeIt2 (pre-phasing), IMPUTE2 (imputation)	ProbABEL	6,563,971	0.997
HNR	[184]	Illumina: Omni Express, Omni1, CoreExomeA and CoreExomeB			$<10^{-5}$		Subject heterozygosity >5 SD away from the mean	MAF ≥0.01 and ≤99.9	First 10 PCs	10	Omni1: 682,618 OmniEx: 646,304 CoreExB: 255,584 CoreExA: 256,445	Impute v.2.3.0	SNPTEST	Excluded due to sample size	
LURIC	[185]	Affymetrix 6.0	Birdseed v.2	≥98%	0.0001	-	-	≥1%	First 3 PCs	3	686,195	IMPUTE v.2	SNPtest v.2.5	7,270,779	1.003

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Study	**R**	**Array**	**Calling Algorithm**	**Per variant call rate**	**HWE p-value**	**Mendelian errors**	**Excess heterozygosity**	**MAF**	**Selection criteria for PCs**	**PCs**	***N* variants for imputation**	**Imputation software**	**GWAS Statistical Analysis**	***N* variants analyzed**	**Inflation factor, lambda**
MDCS	[186]	Illumina Human Omni Express Exome 1.0	GenomeStudio	≥95%	0.0001	-	-	≥1%	All PCs u-ssociated, p>0.05	0	816,728	IMPUTE v.2	SNPtest v.2.5	I: 8,981,701 P: 5,392,317	I: 0.99 P: 1.00
MESA	[187,188]	Affymetrix 6.0	Birdseed v1.33	≥95%	$<10^{-6}$	-	-	≥1%	First 2 PCs	2	881,666	IMPUTE2	ProbABEL	5,340,434	1.027
MGH AF study	[169]	Affymetrix 6.0	Birdseed	≥97%	$<10^{-6}$	-	-	≥1%	-	0	663,637	IMPUTE v2	PLINK v1.07	6,764,173	1.028
MGH CAMP		Infinium HumanCoreExome-24 BeadChips	zCall (GenomeStudio)	≥95%	$\geq 10^{-6}$	-	-	≥1%	PC1-PC10	10	224,343	IMPUTE2	PLINK v1.08	8,262,143	1.01
MGH Stroke	[3,189]	Affymetrix 6.0 + Illumina 610	Birdseed / GenCall	>95% MAF >5%	$<10^{-6}$	-	>±3 SD from the mean	>5%	-	2	GASROS Affymetrix: 579,083 GASROS Illumi-: 398,434 GOCHA: 521,363	IMPUTE2 v.2.3.0	SNPtest v.2.4.1	Excluded due to sample size	
WTCCC2 Munich	[3,190]	Illumina 660	GenCall	>98%	$>10^{-5}$	-	-	>1%	-	0	495,851	MACH+minimac	SNPTEST	5,891,675	1.019

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
PIVUS	[191]	Illumina OmniExpress +Metabochip	GenCall	≥99% (MAF< 5%) or ≥95% (MAF≥ 5%)	$>10^{-6}$	-	>3 SD from the mean	≥1%	First 2 PCs	2	738,879	IMPUTE v.2.2.2	SNPTEST v.2.5	6,045,282	1.006
PREVEND	[192]	Illumina CytoSNP12 v2	GenomeStudio	>95%	$>10^{-6}$	-	-	≥1%	First 5 PCs	5	232,571	IMPUTE1	SNPTEST v.2	5,091,540	1.031
PROSPER	[193]	Illumina Beadchip 660Quad	BeadStudio	≥98%	$<10^{-6}$	-		>1%		4	557,192	IMPUTE v.2.2.2	SNPTEST	7,819,558	1.009
RS	[194]	Illumina Infinium HumanHap550 chip v3.0	BeadStudio	≥98%	$<10^{-6}$	-	>0.336	>1%	First 4 PCs	4	512,849	Mach 1 vs 1.0.151	ProbABEL	RSI: 7,695,631 RS2: 5,543,119 RS3: 5,224,770	P&I RS1: 1.022 RS2: 1.003 RS3: 1.033
SPHFC	[195]	Affymetrix Axion Brazilian Biobank Array	Birdseed v.2	≥97%	$<10^{-6}$	-	-	≥1%	First 3 PCs			IMPUTE v3	PLINK v1.08	7,104,209	1.02
SHIP	[196]	Affymetrix Genome-Wide Human SNP Array 6.0	Birdseed2	≥80%	>0.0001	-	-	≥1%	First 10 PCs	-	905,910	IMPUTE v.2.2.2	QUICKTEST v0.95	5,289,189	0.997
TWINGENE	[197]	Illumina HumanOmniExpress	GenCall	≥97%	$>10^{-7}$	-	>5 SD from the mean	≥1%	First 3 PCs	3	644,556	minimac (release 2012-10-03)	SNPTEST v.2.5	7,201,417	0.983

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Study	**R**	**Array**	**Calling Algorithm**	**Per variant call rate**	**HWE p-value**	**Mendelian errors**	**Excess heterozygosity**	**MAF**	**Selection criteria for PCs**	**PCs**	***N* variants for imputation**	**Imputation software**	**GWAS Statistical Analysis**	***N* variants analyzed**	**Inflation factor, lambda**
ULSAM	[198]	Illumina Omni2.5+Metabochip	GenCall	≥99% (MAF< 5%) or ≥95% (MAF≥ 5%)	$>10^{-6}$	-	>3 SD from the mean	≥1%	First 2 PCs	2	1,587,454	IMPUTE v.2.2.2	SNPTEST v.2.5	7,297,774	0.996
WGHS	[199]	Illumina HumanHap 300 DuoPlus	BeadStudio v. 3.3	≥90%	$>10^{-6}$	-	-	≥1%	PCs 1,2, & 10	3	332,927	MaCH v.1.0.16 + minimac (release 5/29/2012)	ProbABEL	8,144,887	1.02

Supplementary Table 23. General principles for quality control and filtering

Pre-imputation:
Per marker quality control: Call rate (exclude markers if <95%) Hardy-Weinberg Equilibrium (exclude markers if marked deviation) Duplicate concordance (exclude markers with high discordance rates) Mendelian inconsistencies (exclude markers with an excess of Mendelian inconsistencies) Genotype completeness (exclude markers with relatively high missingness) Polymorphism check (exclude monomorphic markers which can represent assay failures) **Per individual quality checks typically include:** Principal Component Analysis Exclude samples with high degree of missingness Exclude samples with unusual heterozygosity Exclude monomorphic markers which can represent assay failures **Exclude related individuals for non-family studies**
Imputation:
Cases and controls imputed together **Criteria for imputation:** 1000G release used for imputation: 20110521 Phase 1 Integrated release ALL Gene reference assembly: GRCh37 SNPs oriented to forward/+ strand
Individuals study analysis:
Account for genotype uncertainty of imputed SNPs **Control for population stratification**
Meta-analysis:
Criteria for including variants (GWAS/EWAS) Imputation quality >0.3 MAF ≥0.01 (GWAS), MAF ≥0.005 (EWAS) Variant present in ≥2 studies Effect allele frequency x imputation quality (INFO) x number of cases ≥10 **Criteria for including genes (gene based tests)** Cumulative MAF per gene ≤0.005 **Quality control:** Estimate genomic inflation factor lambda for each study, and adjust if lambda >1 Check distribution of meta-analysis –log10(p-values) using QQ plots

Supplementary Table 24. Exome chip information per study

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	Total *N* variants analyzed
AFLMU/ MGH AF	[169]	Illumina Infinium HumanExome BeadChip v1.0	CHARGE	-	-	-	Exclude |het|>5 SD	-	p<0.01 in association adjusted for age and sex; derived under exclusion of candidate regions	11	241,465
AGES	[170]	Illumina Exome Chip v1.0	Illumina GenomeStudio2011.1	≥95%	$<10^{-6}$	-	-	-	p<0.05	0	247,501
ARIC	[172]	Illumina HumanExome Beadchip v.1.0	Centrally at CHARGE	0.95	-	-	-	-	First 10 PCs	10	223,577
BBJ	[175]	Infinium OmniExpressExome-8 BeadChip Kit	Illumina GenCall	>0.99	$>10^{-6}$ in control	no trios in samples; QC done using IBS	Yes	Exclude monomorphic in either control or case	Eigenstrates	2	61,024
BEAT-AF	[174]	Illumina HumanCoreExome	BeadStudio	≥95%	$>10^{-6}$	-	> 3 SD from the mean removed	ALL	First 10 PCs	10	495,970
BioMe	[176]	Illumina HumanOmniExpress Exome-8 v1.0	zCall (GenomeStudio)	≥90%	$>10^{-6}$	-	-	≥1%	first 4 PCs	4	241,465
BioVU	[177]	Illumina Infinium HumanExome BeadChip	GenomeStudio	>0.95	$>10^{-6}$	>1 removed	Yes (rate >0.44)	-	first 3 PCs	3	247,039
CHS	[178]	Illumina HumanExome BeadChip v1.0	GenomeStudios	≥97%	None	Any among CEPH trio controls	None	None	5 unless others are associated with the outcome	5	247,870

Supplementary Note - **Novel genetic loci for atrial fibrillation**

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	Total *N* variants analyzed
FHS	[180,181]	Illumina HumanExome BeadChip v1.0	GenomeStudio v. 2011.1 and zCall following CHARGE protocol[200]	-	-	-	-	-	p<0.01 in association adjusted for age and sex	0	247,501
GS:SFHS	[183]	Illumina HumanExome Beadchip v.1-A	GenomeStudio v. 2011.1 CHARGE protocol	0.98	-	-	-	Remove Monomorphic	First 3 PCs	1	247,870
KORA	[201,202]	Illumina Infinium HumanExome BeadChip v1.0	CHARGE	-	-	-	Exclude |het|>5 SD	-	p<0.01 in association adjusted for age and sex; derived under exclusion of candidate regions	11	241,465
LURIC	[185]										Excluded
MESA	[187,200]	Illumina Exome Chip v1.0	GenomeStudio v. 2011.1 and zCall following CHARGE protocol	0.95	$>10^{-6}$	-	-	ALL	Eigenstrates	2	247,039
MGH CAMP		Infinium HumanCoreExome-24 BeadChips	zCall (GenomeStudio)	≥95%	$\geq 10^{-7}$	-	-	≥1%	First 10 PCs	10	247,501
RS	[194]	Illumina Human Exome BeadChip v1.0	zCall following CHARGE	<0.97	-	-	Het excess >0.1 AND Het excess ≤0.9	28,471 monomorphic SNPs were excluded (MAF<1E-9)	First 5	5	247,870
SHIP/SHIP-Trend	[196]	Illumina HumanExome Beadchip v.1.0	SOP v5, zCall v3.3	-	-	-	-	-	First 10 PCs	First 10 PCs	247,039
WGHS	[199,203]	Illumina HumanExome Beadchip v.1.1A	GenomeStudio v. 2011.1 and zCall following CHARGE protocol	0.95	-	-	-	-	-	0	247,727
WHI - CT		Illumina Human	GenomeStudio	0.95	-	-	-	-	Plink	2	246,670

Study	**R**	**Array**	**Calling Algorithm**	**Per variant call rate**	**HWE p-value**	**Mendelian errors**	**Excess heterozygosity**	**MAF**	**Selection criteria for PCs**	**PCs**	**Total *N* variants analyzed**
		Exome BeadChip v1.0	v2010.3								
WHI - OS		Illumina Human Exome BeadChip v1.0	GenomeStudio v2010.3	0.95	-	-	-	-	Plink	2	246,670

Supplementary Table 25. Baseline characteristics of African American ancestry replication studies

	Cases	Controls	Total
N	447	442	889
Women, %	44	48	46
Age at enrollment, mean (SD)	55 (11)	61 (14)	60 (14)
Age at diagnosis, mean (SD)	58 (14)	-	-
Age range (Q1-Q3)	50-61	52-72	51-69
HTN, %	88	87	88
DM, %	37	41	39
HF, %	24	8	16
MI, %	8	3	6

SD, standard deviation; HTN, hypertension; DM, diabetes mellitus; HF, heart failure; MI, myocardial infarction.

Supplementary Table 26. Results from replication in African American ancestry studies

rsID	Risk allele	RAF, %	OR	95% CI	P-value
rs115339321	T	97	1.53	0.82-2.18	0.18
rs79433233	A	3	1.36	0.75-2.47	0.31

RAF, risk allele frequency; OR, odds ratio; CI, confidence interval.

Supplementary Table 27. Results from DEPICT pathway analysis of GWAS meta-analysis results

Original gene set ID	Original gene set description	Nominal P-value
KEGG ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY ARVC	KEGG ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY ARVC	1.27×10^{-6}
KEGG_TIGHT_JUNCTION	KEGG TIGHT JUNCTION	1.75×10^{-6}
MP:0003157	impaired muscle relaxation	2.28×10^{-6}
GO:0016459	myosin complex	8.31×10^{-6}
GO:0060429	epithelium development	1.17×10^{-5}
MP:0000751	myopathy	1.25×10^{-5}
GO:0030855	epithelial cell differentiation	1.67×10^{-5}
KEGG HYPERTROPHIC CARDIOMYOPATHY HCM	KEGG HYPERTROPHIC CARDIOMYOPATHY HCM	3.07×10^{-5}
REACTOME MUSCLE CONTRACTION	REACTOME MUSCLE CONTRACTION	4.18×10^{-5}
GO:0031589	cell-substrate adhesion	8.50×10^{-5}

Supplementary Table 28. Top 5 enriched canonical pathways from Ingenuity Pathway Analysis of GWAS meta-analysis results

Ingenuity Canonical Pathways	P-value	Ratio	Molecules
Coagulation System	0.0088	3/35 (8.6%)	F11, KLKB1, PLAU
Clathrin-mediated Endocytosis Signaling	0.011	7/197 (3.6%)	MET, UBD, FGF17, ACTR2, AAK1, HIP1, PCYOX1
Protein Ubiquitination Pathway	0.013	8/255 (3.1%)	UBD, UBE2G2, USP18, UBE2Q1, BAG1, PSMD5, USP54, PSMD3
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	0.018	2/17 (11.8%)	FDPS, PMVK
Ephrin Receptor Signaling	0.02	6/174 (3.4%)	ACTR2, SHC1, EFNA3, CREB5, EFNA4, EFNA1

Supplementary Table 29. Enriched diseases or functions annotation from Ingenuity canonical pathway analysis of GWAS meta-analysis results

Diseases or Functions Annotation	P-value	*N* molecules	Molecules
Arrhythmia of heart ventricle	$3.0x10^{-9}$	12	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, PKP2, SCN10A, SCN5A, TBX5, THRA, TTN
Ventricular tachycardia	$1.7x10^{-8}$	10	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, PKP2, SCN5A, TBX5, THRA
Tachycardia	$2.5x10^{-8}$	11	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, PITX2, PKP2, SCN5A, TBX5, THRA
Arrhythmia	$5.0x10^{-8}$	16	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, NR3C1, PITX2,PKP2, PLN, SCN10A, SCN5A, TBX5, THRA, TTN, TUBA8
Ventricular fibrillation	$9.5x10^{-7}$	7	DSG2, KCNG2, KCNJ5, PKP2, SCN5A, THRA, TTN
Cardiomyopathy of heart ventricle	$1.2x10^{-6}$	6	CAV1, DSG2, HCN4, PKP2, SCN5A, TTN
Cardiac fibrillation	$1.6x10^{-6}$	11	DSG2, KCNG2, KCNJ5,NR3C1, PITX2,PKP2, PLN, SCN5A, THRA, TTN, TUBA8
Hypertrophy of cardiac muscle	$5.5x10^{-6}$	10	CAV1, CSF3, FBXO32, IL6R, mir-23, PLAU, RAB1A, SHC1, TBX5, TTN
Arrhythmogenic right ventricular dysplasia	$5.7x10^{-6}$	5	DSG2, HCN4, PKP2, SCN5A, TTN

2. Detailed Description of participating studies

The meta-analyses described in this manuscript included the following studies described elsewhere: The **Age, Gene/Environment Susceptibility Study (AGES) Reykjavik study**[169]**,** the **Atrial Fibrillation Biobank LMU (AFLMU)** in the context of the **Arrhythmia-Biobank-LMU** (formerly known as **AFNET**) and the **Cooperative Health Research in the Region of Augsburg (KORA)**[169]**,** the **Atherosclerosis Risk in Communities (ARIC) study**[169]**, Cleveland Clinic Lone Atrial Fibrillation GeneBank Study (CCAF)**[169]**,** the **Cardiovascular Health Study (CHS)**[169], **Framingham Heart Study (FHS)**[169], **Massachusetts General Hospital (MGH) AF study**[169], the **Rotterdam Study (RS)**[169], the **Study of Health in Pomerania (SHIP)**[169], **BioVU**[212], the **Women's Genome Health Study (WGHS)**[169], The **PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)**[175], **Biobank Japan (BBJ)**[175], in addition to the studies described here:

ANGES: The Angiography and Genes Study (ANGES) population consists of 1,000 Finnish individuals participating in the ongoing ANGES study. Angiographic, genetic, and covariate data was available for 808 individuals (516 men and 292 women; mean age 62±10). The data was collected between September 2002 and July 2005. All patients underwent coronary angiography at Tampere University Hospital due to clinically suspected coronary artery disease. The study is a cross-sectional study, and after the angiography, patients were treated according to the Finnish Current Care Guidelines. Patients were also interviewed by a study nurse, and a questionnaire was used to collect general information - age, sex, body mass index, alcohol consumption, smoking, medication, as well as traditional risk factors of atherosclerosis and myocardial infarction. The study has been approved by the Ethics Committee of Pirkanmaa Hospital District and written informed consent was obtained from each patient.

BEAT-AF: The Basel Atrial Fibrillation Cohort Study (BEAT-AF) is a prospective observational, multicenter cohort study. Between 2010 and 2014, 1550 patients with documented atrial fibrillation were enrolled across 7 centers in Switzerland. Exclusion criteria were the inability to sign informed consent and the presence of short transient forms of atrial fibrillation. At baseline, patients completed detailed questionnaires about personal, medical, nutritional and lifestyle factors, current atrial fibrillation symptoms and co-morbidities. Current medications were recorded. A resting 12-lead electrocardiogram (ECG) was recorded and all patients underwent venous blood sampling at the local study center, including DNA from leukocytes. Yearly follow-ups by mailed questionnaires and phone interviews were performed in all patients in order to collect similar information as at baseline and to obtain details about adverse events.
Referents were enrolled from the 'genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors' (GAPP) study, which is an ongoing prospective population-based cohort study among healthy adults in the Principality of Liechtenstein. Between 2010 and 2013, all inhabitants of the Principality of Liechtenstein aged between 25 and 41 years were invited and 2170 agreed to participate in the study. Main exclusion criteria were established cardiovascular disease, chronic kidney disease, diagnosed sleep apnea, a body mass index (BMI) > 35 kg/m2, intake of antidiabetic drugs or any other severe illness. Examinations included detailed assessment of personal, medical, lifestyle and nutritional factors, standardized assessment of weight, height and waist circumference, blood pressure measurement, electrocardiography, bioimpedance analysis, blood, urinary and genetic sampling, spirometry and sleep pulse oximetry with nasal flow measurement. Follow-up examinations are scheduled every 3-5 years. The detailed study design has previously been published.[174]

Bio*Me*: The Mount Sinai Bio*Me* Biobank is an ongoing, prospective, hospital- and outpatient- based population research program operated by The Charles Bronfman Institute for Personalized Medicine (IPM) at Mount Sinai and has enrolled over 33,000 participants since September 2007. Bio*Me* is an

Electronic Medical Record (EMR)-linked biobank that integrates research data and clinical care information for consented patients at The Mount Sinai Medical Center, which serves diverse local communities of upper Manhattan with broad health disparities. Bio*Me* populations include 25% of African ancestry (AA), 36% of Hispanic Latino ancestry (HL), 30% of white European ancestry (EA), and 9% of other ancestry. The BioMe disease burden is reflective of health disparities in the local communities. Bio*Me* operations are fully integrated in clinical care processes, including direct recruitment from clinical sites waiting areas and phlebotomy stations by dedicated recruiters independent of clinical care providers, prior to or following a clinician standard of care visit. Recruitment currently occurs at a broad spectrum of over 30 clinical care sites.

Information on atrial fibrillation, age, sex, body mass index (BMI), type 2 diabetes (T2D), hypertension (HYP), heart failure (HFAIL), and myocardial infarct (MI) was derived from participants' EMRs: Age, sex and BMI were derived from the day of enrolment to the BioMe biobank. Prevalent atrial fibrillation cases were defined as Bio*Me* participants with the ICD-9 code 427.31 (atrial fibrillation) and/or 427.32 (atrial flutter) and controls as individuals who have had ECG's but did not have atrial fibrillation or flutter ICD-9 codes. HYP, HFAIL, and MI were defined using the ICD-9 codes 401.*, 428.*, and 410.*, respectively. In addition to the ICD-9 codes, also individuals taking antihypertensive drugs were considered as having HYP. T2D was defined using the eMerge T2D case and control definition algorithms.[213] The algorithms used were developed by a multidisciplinary team of scientists, clinicians and software specialists and have been validated with excellent performance statistics; 100% sensitivity and >98% positive predictive value for cases, and ≥98% sensitivity and ≥98% positive predictive value for controls.

Bio*Me* participants were genotyped with the Illumina HumanOmniExpressExome-8 v1.0 beadchip array and imputed to the 1000 Genomes Project Phase 1 (March12) reference panel using IMPUTE2. Genome-wide association studies (GWAS) were carried out using SNPTEST 2.4.1 after stratifying by self-reported ancestry (AA: 174 atrial fibrillation cases and 2130 controls; EA: 291 atrial fibrillation cases and 860 controls; HL: 277 atrial fibrillation cases and 3081 controls) and adjustment for a) age, sex and the first 4 GWAS PCs (Model1) and b) age, sex, BMI, T2D, HYP, HFAIL, MI, and the first 4 GWAS PCs (Model2). To ensure high quality of the association results, variants with imputation quality < 0.3, Hardy-Weinberg p-value < $1x10^{-5}$ or minor allele frequency < 0.01 were excluded.

BioVU*:* BioVU is the Vanderbilt University Medical Center's biorepository linked to de-identified electronic health records. BioVU operations[212] and ethical oversight[214] have been described elsewhere. Briefly, DNA is collected from discarded blood samples remaining after routine clinical testing at Vanderbilt outpatient clinics in Nashville, Tennessee and surrounding areas, and is linked to a de-identified version of the patient's electronic health record termed the "Synthetic Derivative." atrial fibrillation cases were defined as individuals who were aged >18 years, had an ICD-9 diagnosis for atrial fibrillation or flutter (ICD-9: 427.3, 427.31, and 427.32), or a cardiologist diagnosis of atrial fibrillation as identified by a natural language processing tool from the unstructured free text of the ECG impression. In all instances, patients with a history of a heart transplant were excluded (Current Procedural Terminology: 33935, 3394, and 580; ICD-9: V42.1, 996.83).[177]

Corogene: The Corogene study was designed as a large cohort to study mainly CAD, but also other related heart diseases such as heart failure and aortic valve disease. We selected the patients from the CAD point of view, and decided to include over 5000 consecutive patients assigned for coronary angiogram. In Finland, coronary angiogram is performed to practically all patients assigned for invasive heart examination. Despite technical developments in diagnostics, coronary angiogram is still the gold standard for evaluating coronaries. The purpose of this study is to follow contemporary trends in coronary heart disease, and related heart disease risk factors, genetics and epigenetics by collecting

cohorts referred to heart examination. New cohorts will be collected at 5-year intervals in order to see trends in CAD, its risk factors and epigenetics.

FINCAVAS: The purpose of the Finnish Cardiovascular Study (FINCAVAS) is to construct a risk profile - using genetic, haemodynamic and electrocardiographic (ECG) markers - of individuals at high risk of cardiovascular diseases, events and deaths. All patients scheduled for an exercise stress test at Tampere University Hospital, who gave informed consent to participate, were recruited between October 2001 and December 2007. The total number of participants was 4,567. In addition to repeated measurements of heart rate and blood pressure, digital high-resolution ECG at 500 Hz was recorded continuously during the entire exercise test, including the resting and recovery phases. About 20% of the patients were examined with coronary angiography. Genetic variations known or suspected to alter cardiovascular function or pathophysiology were analyzed to elucidate the effects and interactions of these candidate genes, exercise, and commonly used cardiovascular medications.

GS:SFHS: Generation Scotland: Scottish Family Health Study (GS:SFHS) is a family-based genetic epidemiology study of ~24,000 volunteers from ~7000 families across Scotland with the capacity for follow-up through record linkage and re-contact. Participants completed a demographic, health and lifestyle questionnaire and provided biological samples including DNA, and ~21,500 participants underwent detailed clinical assessment, including anthropometric, cardiovascular, respiratory, cognition and mental health. Genetic analysis (GWAS) is complete on 20,000 participants with full baseline data and CHI linkage, with linkage to SMR, prescriptions and dental records. A full cohort description can be found elsewhere.[183] Atrial fibrillation was ascertained as a diagnosis of atrial fibrillation by linkage to one or more inpatient visits with ICD-10 code I48 or ICD-9 427.31 in the Scottish Morbidity Record (SMR1) database before or after recruitment to GS:SFHS.

HNR: The study population of the Heinz Nixdorf Recall (HNR) study has been described in detail elsewhere.[184] Approved by the relevant institutional ethics committees, the study follows strict internal and external quality assurance protocols. Briefly, the study cohort comprises 4,814 men and women aged 45 – 75 years from the three adjacent Ruhr cities Essen, Bochum and Mülheim/Ruhr. The vast majority of the study population is of central European ancestry. The study area covers a region of approximately 600 km² with almost 1.2 million inhabitants. Subjects were randomly selected from statutory lists of residence and gave informed consent. The baseline examinations were from 2000-2003, the 5-Year follow-Up from 2006-2008 and the 10-Year follow-up from 2011-2015. A standardized digital 12-lead resting surface ECG was sampled at 250 Hz and recorded on a MAC 5000® ECG recorder (GE Healthcare, Freiburg, Germany). ECGs were interpreted automatically using the integrated 12SL-Code® [12SL ECG analysis with age & gender specific criteria. Physician's guide. PN 416791-004 Revision A. GE Medical Systems IT, 2000]. ECG findings were coded and transferred to our database. The ECG-codes #161 and #162 are for atrial fibrillation and atrial flutter, respectively and were combined for the purpose of this analysis.

LURIC: The Ludwigshafen Risk and Cardiovascular Health (LURIC) study is an ongoing prospective study of more than 3,300 individuals of German ancestry in whom cardiovascular and metabolic phenotypes (CAD, MI, dyslipidemia, hypertension, metabolic syndrome and diabetes mellitus) have been defined or ruled out using standardized methodologies in all study participants.[185] Inclusion criteria for LURIC were: German ancestry (limitation of genetic heterogeneity), clinical stability (except for acute coronary syndromes) and availability of a coronary angiogram. Exclusion criteria were: any acute illness other than acute coronary syndromes, any chronic disease where non-cardiac disease predominated and a history of malignancy within the last five years. Genome-wide analyses using the Affymetrix 6.0 have

been completed in all participants. A 10-year clinical follow-up for total and cause specific mortality has been completed.

MDCS: The Malmö Diet and Cancer study (MDCS) is a community-based prospective epidemiologic cohort of middle-aged individuals from Southern Sweden.[186] In total, 30,447 subjects attended a baseline exam in 1991-1996, when they filled out a questionnaire and underwent anthropometric and blood pressure measurements. Prevalent or incident cases of atrial fibrillation were ascertained from nation-wide hospital registers with high validity as described previously.[186] Genome-wide genotyping of single nucleotide variants was performed using the Illumina Human Omni Express Exome BeadChip kit. Genotyping was performed in a nested case-cohort design, including a random subset of 5878 subjects.

MESA: The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease (disease detected non-invasively before it has produced clinical signs and symptoms) and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. The cohort is a diverse, population-based sample of 6,814 asymptomatic men and women aged 45-84. Approximately 38 percent of the recruited participants are white, 28 percent African American, 22 percent Hispanic, and 12 percent Asian (predominantly of Chinese descent). Participants were recruited during 2000-2002 from 6 field centers across the U.S. (at Wake Forest University; Columbia University; Johns Hopkins University; the University of Minnesota; Northwestern University, and the University of California – Los Angeles). All underwent anthropomorphic measurement and extensive evaluation by questionnaires at baseline, followed by 4 subsequent examinations at intervals of approximately 2-4 years. Age and sex were self-reported. Current atrial fibrillation at baseline was an exclusion criterion. Follow-up phone calls to study participants (every 9-12 months) were used to identify all hospitalizations. Medical records, including discharge diagnoses, were obtained for each hospitalization. Incident atrial fibrillation was defined by International Classification of Disease codes 427.31 or 427.32 (9th revision). In addition, new diagnoses of atrial fibrillation were identified at follow-up by the presence of atrial fibrillation or atrial flutter on a study ECG at Exam 5 (approximately 10 years after baseline). Further information can be found at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000209.v13.p3.

MGH CAMP: The MGH Cardiology and Metabolic Patient (MGH CAMP) cohort comprises 3857 subjects recruited between 2008 and 2012. Two thirds of the subjects were drawn from patients who had appointments with a physician in the MGH Heart Center, whereas one third were recruited independent of any hospital visit. All subjects had plasma and serum samples collected, as well as blood for genomic DNA. Subjects with known diabetes had vascular reactivity measurements (FMD of brachial artery), while subjects without known diabetes had an oral glucose tolerance test. Exome Core Chip genotyping was performed on all subjects. Atrial fibrillation was defined as a self-reported history of fibrillation or flutter at study enrollment, or based on a validated medical record ascertainment algorithm (PPV 88%) that utilizes electrocardiographic and relevant diagnostic, procedure, and medication data.[215]

MGH Stroke study: The Genetics of Cerebral Hemorrhage on Anticoagulation (GOCHA) study is a multicenter study of the genetics of intracerebral hemorrhage in the USA, based at the Massachusetts General Hospital. The cases are individuals presented with acute primary hemorrhagic stroke, aged more than 55 years. The controls were recruited from ambulatory clinics in the same centers in which cases were enrolled.

The Genes Affecting Stroke Risk and Outcome Study (GASROS) is a single-center prospective cohort that enrolled cases with acute ischemic stroke, aged more than 18 years who presented to MGH from 2003 to 2011. Ischemic stroke was defined as a clinical syndrome of associated with a radiographically

proven acute infarction consistent with a vascular pattern and without radiographic evidence of a demyelinating or neoplastic disease or other structural disease. In all subjects, the diagnosis was confirmed by diffusion weighted imaging (DWI) completed within 48 hours after symptom onset. Only patients of self-reported European ancestry were enrolled. Controls were matched to cases on the basis of age, sex and race/ethnicity.

In both GOCHA and GASROS, atrial fibrillation status was determined by reviewing medical records, and/or interview subjects or their families. The diagnosis of atrial fibrillation was established if the subject either had a pre-existing diagnosis or was diagnosed with atrial fibrillation in the hospital. The diagnosis was not confirmed by ECG in all cases.

PIVUS: The participants were randomly sampled from all men and women at age 70 living in Uppsala County in 2001 (www.medsci.uu.se/PIVUS). Of the 2025 individuals invited, 1016 participated. The participants underwent a medical examination including a detailed questionnaire on lifestyle and socioeconomic factors, fasting blood sampling, blood pressure measurement and anthropometric measurements, as previously described.[191] Blood and plasma samples have been frozen until analysis, and blood tests performed include a wide variety of traditional and more recent CVD risk factors, along with DNA extraction. In addition, the individuals have undergone extensive phenotyping including whole body MRI, echocardiography, endothelial function measurements, carotid ultrasound, DXA, and spirometry. The participants have been re-examined at age 75 and 80. Atrial fibrillation was defined by 12-lead ECG at the examinations, as well as diagnosis of atrial fibrillation or flutter in the Swedish National Patient Register before or after the baseline examination (inpatient and specialist outpatient care; ICD-9 code, 427.3 and ICD-10 code, I48).

PREVEND: The PREVEND cohort study was founded in 1997, and is an ongoing community-based cohort study including 8592 inhabitants of the city of Groningen, The Netherlands. PREVEND is investigating the natural course of microalbuminaria and its relation to renal and cardiovascular disease. Details of the protocol, atrial fibrillation ascertainment and covariate definitions have been described elsewhere (www.prevend.org). The baseline screening program consisted of 2 outpatient visits to assess demographic factors, anthropometric measurements, cardiovascular and metabolic risk factors, and health behavior and to collect blood samples and 2 24-h urine samples on 2 consecutive days. Participants were seen at 3-year intervals in the PREVEND outpatient clinic. Atrial fibrillation was ascertained if either atrial flutter or atrial fibrillation was present on a 12-lead ECG obtained at one of the three PREVEND follow-up visits, or at an outpatient visit or hospital admission in the two hospitals in the city of Groningen (University Medical Center Groningen and Martini Hospital). Participants without an electrocardiogram (ECG) (n=248), as well as participants with prevalent atrial fibrillation at the baseline screening (n=79) and without GWAS information (n=4632) were excluded, leaving 3633 for analysis.[216]

SPHFC: Participants for the Sao Paolo Heart Failure Cohort (SPHFC) were prospectively enrolled from the outpatient clinic at the Heart Institute, the University of Sao Paulo Medical School, Sao Paulo, Brazil. Only patients older than 18 years and with symptomatic heart failure (stage C) were enrolled. Different heart failure etiologies were included. Patients with prior myocardial infarction (<3 months), unstable angina, hypertrophic cardiomyopathy, valve heart disease candidates to surgical treatment, obstructive pulmonary disease, severe renal or hepatic dysfunction, current history of cancer, severe peripheral arterial disease, cerebrovascular disease and active infection were excluded. Atrial fibrillation status was determined if either atrial flutter or atrial fibrillation was present on a 12-lead ECG at baseline evaluation or prior and could be confirmed by electronic medical record review.

Supplementary Note - **Novel genetic loci for atrial fibrillation**

TWINGENE: The Swedish Twin Registry contains data regarding health, health-related behaviors, physical activity, eating habits, and environmental stressors, along with other information from Swedish national registries. TWINGENE includes twins born before 1958 that were contacted to participate at the baseline examination between April 2004 and December 2008.[217] Health and medication data were collected from self-reported questionnaires, while blood sampling and in-person testing, including blood pressure measurement and anthropometrics were completed at a local health care center. Several biomarkers, including lipid profiles, fasting glucose, HbA1C and CRP, have been measured, and aliquoted serum is stored at the Karolinska Institutet Biobank. Atrial fibrillation was defined as a diagnosis of atrial fibrillation or flutter in the Swedish National Patient Register before or after the baseline examination (inpatient and specialist outpatient care; ICD-9 code, 427.3 and ICD-10 code, I48).

ULSAM: All men born between 1920 and 1924 in Uppsala, Sweden were invited to participate at age 50 in this longitudinal cohort study that was started in 1970. Participants were reinvestigated at the ages of 60, 70, 77, 82 and 88 years.[198] Blood samples for DNA extraction and main cardiovascular risk factors were available from the investigation at age 70. The participants have undergone extensive phenotyping at repeated time points, including euglycemic clamps, oral glucose tolerance tests, DXA, echocardiography, 24-h ambulatory blood pressure measurement, and a range of biomarkers. Atrial fibrillation was defined by 12-lead ECG at the examinations, as well as diagnosis of atrial fibrillation or flutter in the Swedish National Patient Register (inpatient and specialist outpatient care; ICD-9 code, 427.3 and ICD-10 code, I48).

WHI: The Women's Health Initiative (WHI) is one of the largest (n=161,808) studies of women's health ever undertaken in the United States. The WHI studies consisted of randomized CT, which assigned 68,132 women to active or placebo hormone therapy (HT), dietary modification or control, and/or calcium/vitamin D, supplementation or placebo with specific outcomes of common diseases of aging in women, and also an observational study (OS), which collected data on biological and lifestyle factors and health outcomes. A diverse population including 26,045 (17%) women from minority groups were recruited from 1993-1998 at 40 clinical centers across the U.S. Details of the study design have been previously described.[218,219] For the CT and OS participants enrolled in WHI and who had consented to genetic research, DNA was extracted by the Specimen Processing Laboratory at the Fred Hutchinson Cancer Research Center (FHCRC) using specimens that were collected at the time of enrollment in to the study (between 1993 and 1998).
Baseline atrial fibrillation was determined by an initial questionnaire, which probed for self-reported atrial fibrillation or by presence of atrial fibrillation on the baseline 12-lead electrocardiogram. Women were followed up with a medical history update questionnaire at years 3 to 8, which specifically probed for self-reported atrial fibrillation and hospitalizations.

WTCCC2-Munich: The Wellcome Trust Case Control Consortium 2 Munich (WTCCC2-Munich) study is a hospital-based study on ischemic stroke genetics. Only consecutive European Caucasians recruited from a single dedicated Stroke Unit from South-German origin were selected for this study from the Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-University, Munich. Age, sex and clinical risk factors were collected. Atrial fibrillation was identified by ECG measurement on day of admission. For the German samples controls were Caucasians of German origin participating into the population KORAgen study (www.gsf.de/kora/en/english.html). This survey represents a gender- and age stratified random sample of all German residents of the Augsburg area and consists of individuals 25 to 74 years of age, with about 300 subjects for each 10-year increment. All controls were free of a history of stroke, atrial fibrillation or other cardiovascular diseases.

African American replication studies included:

Penn Medicine Biobank: The Penn Medicine BioBank was started in 2009 and aims to recruit patients within the University of Pennsylvania Health System to donate venous blood. All samples are linked to de-identified electronic medical records. Participation is completely voluntary and written and informed consent are obtained prior to sample collection. For this project, all samples were collected within the inpatient and outpatient sections of the cardiovascular division at the University of Pennsylvania. Atrial fibrillation cases were limited to adults >18 years of age. Atrial fibrillation was ascertained through an ICD-9 diagnosis of atrial fibrillation, atrial flutter or documentation within the medical record.

Duke Biobank: The CATHeterization GENetics (CATHGEN) biorepository collected biospecimens and clinical data on individuals age ≥18 undergoing cardiac catheterization for concern of ischemic heart disease at a single center (Duke University Medical Center) from 2000-2010; a total of N=9334 individuals were collected. Samples were matched at the individual level to clinical data collected at the time of catheterization and stored in the Duke Databank for Cardiovascular Diseases (DDCD). Clinical data included subject demographics, cardiometabolic risk factors, cardiac history including symptoms, age-of-onset of cardiovascular diseases, coronary anatomy and cardiac function at catheterization, laboratory data, and yearly follow-up for hospitalizations, vital status, medication use and lifestyle factors. Atrial fibrillation cases were defined as individuals who had ever had atrial fibrillation based on any ECG available at Duke University or ICD-9 code for atrial fibrillation used for inpatient or outpatient billing.

3. Members of the AFGen Consortium, Neurology working group of the CHARGE Consortium and METASTROKE Consortium

AFGen Consortium Members

Ingrid E. Christophersen, MD, PhD,[1–3] Michiel Rienstra, MD, PhD,[4] Carolina Roselli, MSc,[1,5,6] Xiaoyan Yin, PhD,[7,8] Bastiaan Geelhoed, PhD,[4] John Barnard, PhD,[9] Honghuang Lin, PhD,[7,8] Dan E. Arking, PhD,[10], Albert V. Smith, PhD,[11,12] Christine M. Albert, MD, MPH,[13] Mark Chaffin, MSc[1], Nathan R. Tucker, PhD,[1,2] Molong Li, MD,[2] Derek Klarin, MD,[1] Nathan A Bihlmeyer, BS,[14] Siew-Kee Low, PhD,[15] Peter E. Weeke, MD, PhD,[16,17] Martina Müller-Nurasyid, PhD,[5,18,19] J. Gustav Smith, MD, PhD,[1,20] Jennifer A. Brody, BA,[21] Maartje N. Niemeijer MD,[22] Marcus Dörr, MD,[23,24] Stella Trompet, PhD,[25] Jennifer Huffman, PhD,[26] Stefan Gustafsson, PhD,[27] Claudia Schurman, PhD,[28,29] Marcus E. Kleber, PhD,[30] Leo-Pekka Lyytikäinen, MD,[31] Ilkka Seppälä, MD,[31] Rainer Malik, PhD,[32] Andrea R. V. R. Horimoto, PhD,[33] Marco Perez, MD,[34] Juha Sinisalo, MD, PhD,[35] Stefanie Aeschbacher, MSc,[36,37] Sébastien Thériault, MD, MSc,[38,39] Jie Yao, MS,[40] Farid Radmanesh, MD, MPH,[1,41] Stefan Weiss, PhD,[24,42] Alexander Teumer, PhD,[24,43] Seung Hoan Choi, PhD,[1] Lu-Chen Weng, PhD[1,2] Sebastian Clauss, MD,[2,18] Rajat Deo, MD, MTR,[44] Daniel J. Rader, MD,[44] Svati Shah, MD, MHS,[45] Albert Sun, MD,[45] Jemma C. Hopewell, PhD,[46] Stephanie Debette, MD, PhD,[47–50] Ganesh Chauhan, PhD,[47,48] Qiong Yang, PhD,[51] Bradford B. Worrall, MD, MSc,[52], Guillaume Paré, MD, MSc,[38,39] Yoichiro Kamatani, MD, PhD,[15] Yanick P. Hagemeijer, MSc,[4] Niek Verweij, PhD,[4] Joylene E. Siland, MSc,[4] Michiaki Kubo, MD, PhD,[53] Jonathan D. Smith, PhD,[9] David R. Van Wagoner, PhD,[9] Joshua C. Bis, PhD,[21] Siegfried Perz, MSc,[54] Bruce M. Psaty, MD, PhD,[21,55–57] Paul M. Ridker, MD, MPH,[13] Jared W. Magnani, MD, MSc,[7,58] Tamara B. Harris, MD, MS,[59] Lenore J. Launer, PhD,[59] M. Benjamin Shoemaker, MD, MSCI,[16] Sandosh Padmanabhan, MD,[60] Jeffrey Haessler, MS,[61] Traci M. Bartz, MS,[62] Melanie Waldenberger, PhD,[19,54,63] Peter Lichtner, PhD,[64] Marina Arendt, MSc,[65] Jose E. Krieger, MD, PhD,[33] Mika Kähönen, MD, PhD,[66] Lorenz Risch, MD, MPH,[67] Alfredo J. Mansur, MD, PhD,[68] Annette Peters, PhD,[19,54,69] Blair H. Smith, MD,[70] Lars Lind, MD, PhD,[71] Stuart A. Scott, PhD,[72] Yingchang Lu, MD, PHD,[28,29] Erwin B. Bottinger, MD,[28,73] Jussi Hernesniemi, MD, PhD,[31,74] Cecilia M. Lindgren, PhD,[75] Jorge Wong, MD,[76] Jie Huang, MD, MPH,[77] Markku Eskola, MD, PhD,[74] Andrew P. Morris, PhD,[75,78] Ian Ford, PhD,[79] Alex P. Reiner, MD, MSc,[61,80] Graciela Delgado, Msc,[30] Lin Y. Chen, MD, MS,[81] Yii-Der Ida Chen, PhD,[40] Roopinder K. Sandhu, MD, MPH,[82] Man Li, PhD,[83,84] Eric Boerwinkle, PhD,[85] Lewin Eisele, MD,[65] Lars Lannfelt, MD, PhD,[86] Natalia Rost, MD, MPH, FAAN,[1,87] Christopher D. Anderson, MD, MMSc,[1,41] Kent D. Taylor, PhD,[40] Archie Campbell, MA,[88] Patrik K. Magnusson, PhD,[89] David Porteous, PhD,[88] Lynne J. Hocking, PhD,[90] Efthymia Vlachopoulou, PhD,[91] Nancy L. Pedersen, MA, PhD,[89] Kjell Nikus, MD, PhD,[74] Marju Orho-Melander, PhD,[92] Anders Hamsten, MD, PhD,[93] Jan Heeringa, MD, PhD,[22] Joshua C. Denny, MD,[16] Jennifer Kriebel, PhD,[54,63,69] Dawood Darbar, MD,[94] Christopher Newton-Cheh, MD, MPH,[1,2] Christian Shaffer, BS,[16] Peter W. Macfarlane, PhD, DSc,[95] Stefanie Heilmann, PhD,[96,97] Peter Almgren, MSc,[92] Paul L. Huang, MD, PhD,[2] Nona Sotoodehnia, MD, MPH,[98] Elsayed Z. Soliman, MD, MSc, MS,[99] Andre G. Uitterlinden, PhD,[100] PhD, Albert Hofman, MD, PhD,[22] Oscar H. Franco, MD, PhD,[22] Uwe Völker,

PhD,[24,42] Karl-Heinz Jöckel, PhD,[65] Moritz F. Sinner, MD, MPH,[18,19] Henry J. Lin, MD,[40] Xiuqing Guo, PhD,[40] METASTROKE Consortium of the ISGC, Neurology Working Group of the CHARGE Consortium, Martin Dichgans, MD,[32,101,102], Erik Ingelsson, MD, PhD,[27,103] Charles Kooperberg, PhD,[61] Olle Melander, MD, PhD,[104] Ruth J. F. Loos, PhD,[28,29,105] Jari Laurikka, MD, PhD,[106] David Conen, MD, MPH,[36–38] Jonathan Rosand, MD, MSc,[1,41] Pim van der Harst, MD, PhD,[4] Marja-Liisa Lokki, PhD,[91] Sekar Kathiresan, MD,[1] Alexandre Pereira, MD, PhD,[107] J. Wouter Jukema, MD, PhD,[25,108,109] Caroline Hayward, PhD,[26] Jerome I. Rotter, MD,[110] Winfried März, MD,[111] Terho Lehtimäki, MD, PhD,[31] Bruno H. Stricker, MD, PhD,[112] Mina K. Chung, MD, PhD,[9] Stephan B. Felix, MD,[23,24] Vilmundur Gudnason, MD, PhD,[11,12] Alvaro Alonso, MD, PhD,[113] Dan M. Roden, MD,[16] Stefan Kääb, MD, PhD,[18,19] Daniel I. Chasman, PhD,[1,114] Susan R. Heckbert, MD, PhD,[55,56] Emelia J. Benjamin, MD, ScM,[7,58,115] Toshihiro Tanaka, MD, PhD,[116,117] Kathryn L. Lunetta, PhD,[7,8] Steven A. Lubitz, MD, MPH,[1,2,118] Patrick T. Ellinor, MD, PhD,[1,2,118]

AFGen Consortium Member Affiliations

1. Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA.
2. Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA.
3. Department of Medical Research, Bærum Hospital, Vestre Viken Hospital Trust, Norway.
4. Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.
5. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
6. Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.
7. NHLBI and Boston University's Framingham Heart Study, Framingham, MA, USA.
8. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.
9. Departments of Cardiovascular Medicine, Cellular and Molecular Medicine, Molecular Cardiology, and Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, USA.
10. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
11. Icelandic Heart Association, Kopavogur, Iceland.
12. Faculty of Medicine, University of Iceland, Reykavik, Iceland.
13. Divisions of Preventive and Cardiovascular Medicine, Brigham and Women's Hospital & Harvard Medical School, Boston, MA, USA.
14. Predoctoral Training Program in Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
15. Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
16. Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA.
17. The Heart Centre, Department of Cardiology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark.
18. Department of Medicine I, University Hospital Munich, Ludwig-Maximilians-University, Munich, Germany.

19. DZHK (German Centre for Cardiovascular Research), partner site: Munich Heart Alliance, Munich, Germany.
20. Molecular Epidemiology and Cardiology, Clinical Sciences, Lund University, Lund, Sweden.
21. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA.
22. Department of Epidemiology, Erasmus University Medical Center Rotterdam, Rotterdam, the Netherlands.
23. Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany.
24. DZHK (German Centre for Cardiovascular Research), partner site: Greifswald, Germany.
25. Department of Cardiology, Leiden University Medical Center, The Netherlands.
26. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, UK.
27. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.
28. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
29. The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
30. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Germany.
31. Department of Clinical Chemistry, Fimlab Laboratories and University of Tampere School of Medicine, Tampere, Finland.
32. Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians University, München, Germany.
33. Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil.
34. Stanford University, Stanford, CA, USA.
35. Heart and Lung Center HUS, Helsinki University Central Hospital, Helsinki, Finland.
36. University Hospital Basel, Switzerland.
37. Cardiovascular Research Institute Basel, Switzerland.
38. Population Health Research Institute, Hamilton, Canada.
39. Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Canada.
40. Institute for Translational Genomics and Population Sciences, Department of Pediatrics, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
41. Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA.
42. Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany.
43. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany.
44. Division of Cardiovascular Medicine, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA.
45. Division of Cardiology, Department of Medicine, Duke University School of Medicine, Durham, NC, USA.
46. CTSU - Nuffield Department of Population Health, University of Oxford, Oxford, UK.
47. Inserm Center U1219 (Bordeaux Population Health Centre), Bordeaux, France.
48. University of Bordeaux, Bordeaux, France.

49. Department of Neurology, Bordeaux University Hospital, Bordeaux, France.
50. Department of Neurology, Boston University School of Medicine, Boston, MA, USA.
51. Biostatistics Department, School of Public Health, Boston University, Boston, MA, USA.
52. University of Virginia Health System, Departments of Neurology and Public Health Science, Charlottesville, VA, USA.
53. RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
54. Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
55. Department of Epidemiology and Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA.
56. Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA.
57. Department of Health Services, University of Washington, Seattle, WA, USA.
58. Department of Medicine, Boston University School of Medicine, Boston, MA, USA.
59. Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, MD, USA.
60. Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, UK.
61. Fred Hutchinson Cancer Research Center, University of Washington, Seattle, WA, USA.
62. Cardiovascular Health Research Unit, Departments of Medicine and Biostatistics, University of Washington, Seattle, WA, USA.
63. Research unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
64. Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
65. Institute for Medical Informatics, Biometry, and Epidemiology, University Hospital, University Duisburg-Essen, Germany.
66. Department of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
67. University Institute of Clinical Chemistry, University of Bern, Switzerland and labormedizinisches zentrum Dr. Risch, Schaan, Liechtenstein.
68. Heart Institute, University of Sao Paulo, Sao Paulo, Brazil.
69. German Center for Diabetes Research, Neuherberg, Germany.
70. Division of Population Health Sciences, University of Dundee, Scotland, UK.
71. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden.
72. Department of Genetics and Genomic Sciences , Icahn School of Medicine at Mount Sinai, New York, NY, USA.
73. Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
74. Department of Cardiology, Heart Hospital, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
75. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
76. Division of Cardiology, Hamilton Health Sciences, McMaster University, Hamilton, Ontario, Canada.

77. Boston VA Research Institute, Inc., Boston, MA, USA.
78. Department of Biostatistics, University of Liverpool, Liverpool, UK.
79. Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK.
80. Department of Epidemiology, University of Washington, Seattle, WA, USA.
81. Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, MN, USA.
82. Division of Cardiology, University of Alberta, Edmonton, Canada.
83. Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA.
84. Division of Nephrology & Hypertension, Internal Medicine, School of Medicine, University of Utah, UT, USA.
85. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA.
86. Department of Public Health and Caring Sciences, Geriatrics, Uppsala University, Uppsala, Sweden.
87. Acute Stroke Services, Massachusetts General Hospital, Boston, MA, USA.
88. Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, UK.
89. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
90. Musculoskeletal Research Programme, Division of Applied Medicine, University of Aberdeen, Aberdeen, UK.
91. Transplantation Laboratory, Medicum, University of Helsinki, Helsinki, Finland.
92. Department of Clinical Sciences, Lund University, Malmö, Sweden.
93. Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.
94. University of Illinois, Chicago, IL, USA.
95. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK.
96. Institute of Human Genetics, University of Bonn, Germany.
97. Department of Genomics, Life & Brain Research Center, University of Bonn, Germany.
98. Cardiovascular Health Research Unit, University of Washington Medical Center, Seattle, WA, USA.
99. Epidemiological Cardiology Research Center (EPICARE), Wake Forest School of Medicine, Winston Salem, NC, USA.
100. Department of Epidemiology and Internal Medicine, Erasmus University Medical Center Rotterdam, the Netherlands.
101. Munich Cluster for Systems Neurology (SyNergy), München, Germany.
102. German Center for Neurodegenerative Diseases (DZNE), Munich, Germany.
103. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA.
104. Department of Internal Medicine, Clinical Sciences, Lund University, Malmö, Sweden.
105. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
106. Department of Cardio-Thoracic Surgery, Heart Hospital, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.

107. Laboratory of Genetics and Molecular Biology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil and Department of Genetics, Harvard Medical School, Boston, MA, USA.
108. Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands.
109. Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands.
110. Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
111. Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria and Synlab Academy, Synlab Services GmbH, Mannheim, Germany.
112. Department of Epidemiology and Internal Medicine, Erasmus University Medical Center Rotterdam, the Netherlands and Inspectorate of Health Care, Utrecht, the Netherlands.
113. Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA.
114. Divisions of Preventive Medicine and Genetics, Brigham and Women's Hospital & Harvard Medical School, Boston, MA, USA.
115. Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.
116. Laboratory for Cardiovascular Diseases, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
117. Department of Human Genetics and Disease Diversity, Tokyo Medical and Dental University Graduate School of Medical and Dental Sciences, Tokyo, Japan.
118. Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA.

Neurology Working Group of the CHARGE Consortium Members

Ganesh Chauhan, PhD; Corey R Arnold, BSc; Audrey Y Chu, PhD; Myriam Fornage, PhD; Azadeh Reyahi, MSc; Joshua C Bis, PhD; Aki S Havulinna, DSc; Muralidharan Sargurupremraj, PhD; Albert V Smith, PhD; Hieab H H Adams, MSc; Seung Hoan Choi, MA; Stella Trompet, PhD; Melissa E Garcia, MPH; Alexander Teumer, PhD; Céline Bellenguez, PhD; Xueqiu Jian, PhD; Claudia L Satizabal, PhD; Yongmei Liu, PhD; Susan R Heckbert, MD; Kenneth Rice, PhD; Nicholas L Smith, PhD; Reinhold Schmidt, MD; Hugo J Aparicio, MD; Didier Leys, MD; Claudine Berr, MD; Jean-François Dartigues, MD; Cecilia M Lindgren, PhD; Albert Hofman, MD; Peter J Koudstaal, MD; André G Uitterlinden, PhD; Anton J M de Craen, PhD; Andrew D Johnson, PhD; Hans Jörgen Grabe, MD; Yoichiro Kamatani, MD; Oscar L Lopez, MD; Jerome I Rotter, MD; Rebecca F Gottesman, MD; David S Knopman, MD; B Gwen Windham, MD; Alexa Beiser, PhD; Curtis R French, PhD; Mark Lathrop, PhD; Vilmundur Gudnason, MD; Tobias Kurth, MD; Bruce M Psaty, MD; Tamara B Harris, MD; Stephen S Rich, PhD; Anita L deStefano, PhD; Carsten O Schmidt, PhD; Veikko Salomaa, MD; Thomas H Mosley, PhD; Erik Ingelsson, MD; Cornelia M van Duijn, PhD; Ordan J Lehmann, MD; Christophe Tzourio, MD; Lenore J Launer, PhD; M Arfan Ikram, MD; Peter Carlsson, PhD; Daniel I Chasman, PhD; Sarah J Childs, PhD; William T Longstreth, Jr, MD; Sudha Seshadri, MD; Stéphanie Debette, MD

Neurology Working Group of the CHARGE Consortium member affiliations

University Medicine Greifswald, Institute for Community Medicine, SHIP-KEF, Greifswald, Germany (AT, COS); University Medicine Greifswald, Department of Psychiatry and Psychotherapy, Greifswald, Germany (HJG); Division of Preventive Medicine, Brigham and

Women's Hospital, Boston, MA, USA (DIC, AYC, TK); Harvard Medical School, Boston, MA, USA (DIC); Inserm Research Center for Epidemiology and Biostatistics (U897) - Team Neuroepidemiology, Bordeaux, France (TK, MS, CT, JFD, GC, SD); University of Bordeaux, College of Health Sciences, Bordeaux, France (TK); University of Virginia, Charlottesville, VA USA (SSR); National institute for Health and Welfare, Helsinki, Finland (ASH, VS); Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands (HHHA, AH, CMvD, MAI); Department of Neurology, Erasmus MC, Rotterdam, The Netherlands (PJK); Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands (AGU); Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA (JCB, BMP); Department of Epidemiology, University of Washington, Seattle, WA, USA (BMP, WTL, NLS, SRH); Department of Health Services, University of Washington, Seattle, WA, USA (BMP); Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA (BMP); Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA (JIR); Division of Genomic Outcomes, Department of Pediatrics, Harbor-UCLA Medical Center, California, USA (JIR); UCLA, California, USA (JIR); Department of Biostatistics, University of Washington, Seattle, WA, USA (KR); Department of Neurology; University of Washington, Seattle, WA, USA (WTL); Seattle Epidemiologic Research and Information Center of the Department of Veterans Affairs Office of Research and Development, Seattle, WA, USA (NLS); Department of Cardiology, Leiden University Medical Center, Leiden The Netherlands (ST); Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands (ST, AJMdC); Center for Human Genetics, Division of Public Health Sciences, Wake Forest School of Medicine, Winston Salem, NC, USA (YL); Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Bethesda, MD, USA (TBH, MEG); Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden (EI); Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK (CML); University of Bordeaux, Bordeaux, France (MS, CT, JFD, GC, SD, TK); Université of Montpellier, France (CeB); Institut Pasteur de Lille, Lille, France (CeB); Université Lille Nord de France, Lille, France (DL); Department of Neurology, Lille University Hospital, Lille, France (DL); INSERM U1171, Lille, France (DL); INSERM U1061, Montpellier, France (ClB); University Montpellier 1, Montpellier, France (ClB); Department of Neurology, Bordeaux University Hospital, Bordeaux, France (SD, JFD); Boston University School of Medicine, Boston, MA, USA (SD); IRP/NIA/NIH, Bethesda, MD, USA (LJL); Department of Neurology, University of Pittsburgh, PA, USA (OLL); Icelandic Heart Association, Kópavogur, Iceland (VG, AVS); University of Iceland, Faculty of Medicine, Reykjavik, Iceland (VG); Institute of Molecular Medicine and Human Genetics Center; University of Texas Health Science Center at Houston, Houston, TX, USA (MF, XJ); Department of Neurology; Johns Hopkins University, Baltimore, MD, USA (RFG); Department of Neurology; Mayo Clinic, Rochester, MN, USA (DSK); Department of Medicine, Division of Geriatrics; University of Mississippi Medical Center, Jackson, MS, USA (THM, BGW); McGill University and Genome Quebec Innovation Center (ML); Department of Neurology, Clinical Division of Neurogeriatrics, Medical University Graz, Austria (RS); National Heart, Lung, and Blood Institute‘s Framingham Heart Study Cardiovascular Epidemiology and Human Genomics Branch, Framingham, Massachusetts, USA (ADJ); Fondation Jean Dausset, Centre d‘Etude du Polymorphisme Humain (CEPH), Paris, France (YK); Center for

Integrative Medical Sciences, RIKEN, Yokohama, Japan (YK); Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA (AB, ALD, HJA, SS, CLS, SHC); Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA (AB, ALD, SS, SHC); The National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA (AB, ADJ, ALD, AYC, HJA, SS, CLS, SHC); Department of Ophthalmology, University of Alberta, Edmonton, Alberta, Canada (CRF, OJL); Alberta Children's Hospital Research Institute and Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, Alberta, Canada (CRA, SJC); Department of Medical Genetics, University of Alberta, Edmonton, Alberta, Canada (OJL); Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden (AR, PC).

METASTROKE Consortium Contributors

Study	**Investigator**
ASGC	Elizabeth Holliday
	Chris Levi
BRAINS	Pankaj Sharma
	Iona Cotlarciuc
GASROS	Jonathan Rosand
GEOS	Steven Kittner
	John Cole
HPS	Jemma Hopewell
	Robert Clarke
ISGS/SWISS	James Meschia
	Mike Nalls
MILANO	Giorgio Boncoraglio
VISP	Bradford B. Worrall
	Michéle Sale
	Wei-Min Chen
WHI	Alex Reiner
WTCCC2-D	Martin Dichgans
	Rainer Malik
WTCCC2-UK	Hugh Markus
	Steve Bevan
CADISP	Stephanie Debette
	Ganesh Chauhan
Krakow	Agnieszka Slowik
	Joanna Pera
Leuven	Vincent Thijs
	Robin Lemmens

Lund	Arne Lindgren
Münster	E Bernd Ringelstein
	Klaus Berger
SIGN	Brackie Mitchell
Portugal	Sofia A Oliveira
	Jose M Ferro
	Astrid M Vicente
RACE	Danish Salaheen
USA	Dan Woo

Other METASTROKE participants

ARIC	Myriam Fornage
CHS	Bruce Psaty
	Josh Bis
deCODE	Solveig Gretarsdottir
FHS	Sudha Seshadri
	Qiong Yang
	Claudia L. Satizabal Barrera
HVH	Bruce Psaty
Rotterdam	Arfan Ikram
Barcelona	Israel Fernandez
BSS	Massimo Pandolfo
	Sherine Abboud
Copenhagen	Marianne Benn
	Børge G Nordestgaard
ESS	Cathie Sudlow
	Kristina Rannikmae
Glasgow	Matthew Walters
	Peter Higgins
GoDarts	Colin N A Palmer
	Alexander S F Doney
Graz	Reinhold Schmidt
INTERSTROKE	Guillame Pare
Karolinska	Konstantinos Kostulas
SMART	Paul de Bakker
	Ale Algra
Athero-Express	Sander van der Laan
	Gerald Pasterkamp
Sifap	Anne-Katrin Giese

4. SUPPLEMENTARY RESULTS

Ancestry-specific GWAS meta-analyses

Separate GWAS in 15,993 cases and 113,719 referents of European ancestry revealed one additional association on chromosome 15q21 (rs2921421, OR 1.72, 95% CI 1.42-2.09, P=3.29x10^{-8}, **Supplementary Table 6**); however, there was only one significant variant at this locus and the variant was imputed with low quality across all studies reducing our confidence in this finding. Additional replication in another European ancestry study is needed to clarify the relevance of rs2921421. In meta-analysis of 837 cases and 2456 referents of Asian ancestry we identified an association on chromosome 12q15 (rs7138621, OR 7.92, 95% CI 4.26-14.73, P=6.48x10^{-11}), which was not significant in *in silico* replication in 8180 cases and 28,612 referents in the Biobank Japan (**Supplementary Table 10**). Separate meta-analyses in individuals of Brazilian and Hispanic descent did not identify additional loci; however, our power was limited in each of these sub-groups.

GWAS meta-analyses of incident and prevalent atrial fibrillation in Europeans

Separate GWAS meta-analyses of incident (7232 cases) and prevalent (8656 cases) atrial fibrillation in Europeans showed similar results to the European ancestry analysis (**Supplementary Tables 8-9, Supplementary Figs. 7-8**); however, we did reveal a novel atrial fibrillation locus associated with prevalent atrial fibrillation at chromosome 12p11 (rs1454934, OR 1.16, 95% CI 1.1-1.22, P=4.18x10^{-8}). The most significant variant at this locus was intronic to the gene plakophilin-2 (*PKP2*), which encodes an important component of the desmosome and is known to be associated with arrhythmogenic right ventricular cardiomyopathy[220] and Brugada syndrome.[221,222]

Replication of genetic variants specific to African American ancestry GWAS meta-analysis

The variants rs115339321 (OR 1.53, 95% CI 0.82-2.18, P=0.18) and rs79433233 (OR 1.36, 95% CI 0.75-2.47, P=0.31) were not significantly associated with atrial fibrillation in 447 atrial fibrillation cases and 442 referents of African American ancestry (**Supplementary Table 25-26**). The lack of replication may be caused by the small sample size of the replication study. Further replication in a larger sample of African American ancestry is needed to clarify the role of the variants rs115339321 and rs79433233.

Pathway analyses

1. DEPICT

The most significant pathway identified using the DEPICT software was the arrhythmogenic right ventricular cardiomyopathy (ARVC) pathway (P=1.3x10^{-6}, **Supplementary Table 27**). None of the pathways analyzed reached an FDR <5%.

2. IPA

The most significantly enriched biological pathway was the coagulation system (P=0.0088). In addition, many genes were involved in the clathrin-mediated endocytosis signaling pathway (P=0.011) and the protein ubiquitination pathway (P=0.013). The most significant pathways are listed in **Supplementary Table 28** None of the pathways reached the significance threshold (FDR<5%). In addition, many of the genes investigated were involved in arrhythmia mechanisms (**Supplementary Table 29**).

5. ACKNOWLEDGMENTS

AFLMU/KORA: This work is funded by the European Commission's 7th Framework Programme FP7-HEALTH-2013 No. 602299: EU-CERT-ICD to Dr. Kääb and the European Commission's Horizon 2020 Framework Programme EU-H2020-PHC-RIA No.633196: CATCH-ME to Dr. Sinner. It is further supported by the DZHK (German Centre for Cardiovascular Research), partner site: Munich Heart Alliance, Munich, Germany, the BMBF Spitzencluster personalized medicine m4 (01 EX1021E), the LMU Excellence Initiative (42595-6), and the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ, all to Dr. Kääb. The KORA platform is funded by the BMBF and by the State of Bavaria.

AGES: The Age, Gene/Environment Susceptibility Reykjavik Study has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

ANGES: The Angiography and Genes Study (ANGES) has been financially supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9M048 and 9N035), the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, Finland, and the Tampere Tuberculosis Foundation.

ARIC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. Funding support for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419).

BEAT-AF: Swiss National Science Foundation (PP00P3_133681 and PP00P3_159322); Swiss Heart Foundation; University of Basel; University Hospital Basel

Biobank Japan: The BioBank Japan Project was supported by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government.

Bio*Me*: The Mount Sinai Bio*Me* Biobank Program is supported by The Andrea and Charles Bronfman Philanthropies. Analyses of Bio*Me* data was supported in part through the computational resources and staff expertise provided by the Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

BioVU: The dataset used in the analyses described were obtained from Vanderbilt University Medical Centers BioVU which is supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. Genome-wide genotyping was funded by NIH grants RC2GM092618 from NIGMS/OD and U01HG004603 from NHGRI/NIGMS.

Supplementary Note - **Novel genetic loci for atrial fibrillation**

CCAF is funded by National Institutes of Health grants R01 HL090620 and R01 HL111314 to MKC, JB, JS, and DVW, the NIH National Center for Research Resources for Case Western Reserve University and The Cleveland Clinic Clinical and Translational Science Award UL1-RR024989, and the Department of Cardiovascular Medicine philanthropic research fund, Heart and Vascular Institute, Cleveland Clinic

CHS: This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268200960009C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, R01HL130114, R01HL068986, and R01HL085251 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Corogene: The Corogene study has been financially supported by the Aarno Koskelo Foundation (partial); the Finnish Foundation for Cardiovascular Research (partial); the Paulo Foundation (partial); Jenny and Antti Wihuri Foundation (partial); EVO funds of Helsinki University Central Hospital; Wellcome Trust, UK (the genetic part).

FINCAVAS: The Finnish Cardiovascular Study (FINCAVAS) has been financially supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9M048 and 9N035), the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, Finland, and the Tampere Tuberculosis Foundation.

FHS: This research was conducted using data and resources from Framingham Heart Study (FHS) of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine based on analyses by Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No.N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. Other support came from 1R01 HL092577, 1RO1 HL076784, 1R01 AG028321, Evans Center for Interdisciplinary Biomedical Research ARC on Atrial Fibrillation at Boston University (http://www.bumc.bu.edu/evanscenteribr/the-arcs/the-arcs/ (Benjamin), 6R01-NS 17950 and American Heart Association 09FTF2190028.

GS:SFHS: Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Dirorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award “STratifying Resilience and Depression Longitudinally” (STRADL) (Reference 104036/Z/14/Z).

Supplementary Note - **Novel genetic loci for atrial fibrillation**

HNR: The Heinz Nixdorf Recall Study thanks the Heinz Nixdorf Foundation (Germany) and projects SI 236/8-1 and SI 236/9-1 from the German Research Council for the generous support of this study. We acknowledge the support of the Sarstedt AG & Co. (Nümbrecht, Germany) for laboratory equipment. We are indebted to all study participants and to the dedicated personnel of the study center of the Heinz Nixdorf Recall study. Advisory Board: Meinertz T, Hamburg, Germany (Chair); Bode C, Freiburg, Germany; de Feyter PJ, Rotterdam, Netherlands; Güntert B, Hall i.T., Austria; Gutzwiller F, Bern, Switzerland; Heinen H, Bonn, Germany; Hess O, Bern, Switzerland; Klein B, Essen, Germany; Löwel H, Neuherberg, Germany; Reiser M, Munich, Germany; Schwaiger M, Munich, Germany; Steinmüller C, Bonn, Germany; Theorell T, Stockholm, Sweden; Willich SN, Berlin, Germany.

LURIC: LURIC was supported by the 7th Framework Program (integrated project AtheroRemo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739) of the European Union, by the INTERREG IV Oberrhein Program (Project A28, Genetic mechanisms of cardiovascular diseases) with support from the European Regional Development Fund (ERDF) and the Wissenschaftsoffensive TMO

MDCS: The MDCS was made possible by grants from the Malmö city council. J. Gustav Smith was supported by The Märta Winkler foundation, Swedish Heart Association, Swedish Heart-Lung Foundation, the European Research Council, the Swedish Research Council, the Crafoord Foundation, governmental funding of clinical research within the Swedish National Health Service, Skåne University Hospital in Lund.

MESA: The Multi-Ethnic Study of Atherosclerosis (MESA) is supported by NIH contracts HHSN2682015000031, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and by grants UL1-TR-000040, UL1-TR-001079, and UL1-RR-025005 from NCRR. Funding for MESA Family was provided by grants R01-HL-071205, R01-HL-071051, R01-HL-071250, R01-HL-071251, R01-HL-071252, R01-HL-071258, and R01-HL-071259, and by UL1-RR-025005 and UL1RR033176 from NCRR. Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

MGH AF: This work was supported by grants from the National Institutes of Health to Dr. Ellinor (1RO1HL092577, R01HL128914, K24HL105780). Dr. Christophersen is supported by a mobility grant from the Research Council of Norway [240149/F20]. Dr. Ellinor is also supported by an Established Investigator Award from the American Heart Association (13EIA14220013) and by the Fondation Leducq (14CVD01). Dr. Lubitz is supported by grants from the NIH (K23HL114724) and by a Doris Duke Charitable Foundation Clinical Scientist Development Award (2014105).

MGH CAMP: The recruitment, collection of samples, and genotyping was supported by Pfizer. Analysis of data was a three way collaboration between MGH, the Broad Institute, and Pfizer. Dr. Huang is supported by grants from the NIH (NS33335, NS055104).

MGH Stroke study: The GOCHA study was supported by the grant R01NS059727 from the NINDS. The GASROS study was supported by the grant U01 NS069208-01 from the NINDS.

Supplementary Note - **Novel genetic loci for atrial fibrillation**

PIVUS: These projects were supported by Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), Swedish Diabetes Foundation (2013-024), Swedish Research Council (2012-1397 and 2015-02907), and Swedish Heart-Lung Foundation (20140422). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping was funded by the Wellcome Trust under awards WT064890 and WT086596. Analysis of genetic data was funded by the Wellcome Trust under awards WT098017 and WT090532. Andrew P Morris is a Wellcome Trust Senior Fellow in Basic Biomedical Science, under award WT098017.

PREVEND: The PREVEND study is supported by the Dutch Kidney Foundation (grant E0.13) and the Netherlands Heart Foundation (grant NHS2010B280. Dr. M. Rienstra is supported by grants from the Netherlands Organization for Scientific Research (Veni grant 016.136.055), the EHRA Academic Fellowship program, and from the Netherlands Cardiovascular Research Initiative: an initiative supported by the Netherlands Heart Foundation, CVON 2014–9: "Reappraisal of Atrial Fibrillation: interaction between hyperCoagulability, Electrical remodelling, and Vascular destabilization in the progression of atrial fibrillation (RACE V)".

PROSPER: The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810).

RS: The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam; The Netherlands Organization for Scientific Research; The Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; The Netherlands Heart Foundation; the Ministry of Education, Culture and Science; the Ministry of Health Welfare and Sports; the European Commission; and the Municipality of Rotterdam. Support for genotyping was provided by The Netherlands Organization for Scientific Research (NWO) (175.010.2005.011, 911.03.012) and Research Institute for Diseases in the Elderly (RIDE). This study was supported by The Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) project nr. 050-060-810.

SPHFC: The Heart Failure Cohort was supported by the Samaritan Hospital of São Paulo and Brazilian Ministry of Health - Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde (PROADI-SUS), the Sao Paulo Research Agency Fapesp (grant 2013/17368-0).

SHIP: SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG.

Supplementary Note - **Novel genetic loci for atrial fibrillation**

TWINGENE: These projects were supported by Ministry for Higher Education, GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH grant DK U01-066134, Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), Swedish Diabetes Foundation (2013-024), Swedish Research Council (M-2005-1112, 2009-2298, 2012-1397 and 2015-02907), and Swedish Heart-Lung Foundation (20140422). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. We thank the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se) for excellent genotyping.

ULSAM: These projects were supported by Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), Swedish Diabetes Foundation (2013-024), Swedish Research Council (2012-1397 and 2015-02907), and Swedish Heart-Lung Foundation (20140422). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping was funded by the Wellcome Trust under awards WT064890 and WT086596. Analysis of genetic data was funded by the Wellcome Trust under awards WT098017 and WT090532. We thank the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se) for excellent genotyping.

WGHS: The Women's Genome Health Study (WGHS) is supported by the National Heart, Lung, and Blood Institute (HL043851 and HL080467 [both Buring]) and the National Cancer Institute (CA047988 [Buring] and UM1CA182913 [Buring and Lee]) with collaborative scientific support and funding for genotyping provided by Amgen (Chasman and Ridker). Atrial fibrillation endpoint confirmation was supported by HL-093613 (Albert) and HL116690 (Albert) and a grant from the Harris Family and Watkin's Foundation (Tedrow).

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. A full listing of the WHI investigators can be found at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf. We thank the Women's health Initiative investigators, staff, and study participants for their contributions.

WTCCC2-Munich: This work was supported by grants received from the German Federal Ministry of Education and Research (BMBF) in the context of the e:Med program (e:AtheroSysMed), the FP7 European Union project CVgenes@target (261123), the DFG as part of the CRC 1123 (B3), the Corona Foundation, and the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain).

Penn Biobank: This work was supported by grants from the National Institutes of Health to Dr. Deo (K23-DK089118).

Duke Biobank: This work was supported by grants from the National Heart, Lung and Blood Institute (HL 095987 [Shah] and HL101621 [Kraus].

Role of the Sponsor:
None of the funding agencies had any role in the study design, data collection or analysis, interpretation of the data, writing of the manuscript, or in the decision to submit the manuscript for publication.

6. SUPPLEMENTARY REFERENCES

1. Pruim, R. J. *et al.* LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26,** 2336–2337 (2010).
2. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42,** D1001-6 (2014).
3. Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol.* **11,** 951–62 (2012).
4. Dichgans, M. *et al.* Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke.* **45,** 24–36 (2014).
5. Gretarsdottir, S. *et al.* Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann. Neurol.* **64,** 402–9 (2008).
6. den Hoed, M. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45,** 621–31 (2013).
7. Arking, D. E. *et al.* Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat. Genet.* **46,** 826–36 (2014).
8. Smith, J. G. *et al.* Impact of ancestry and common genetic variants on QT interval in African Americans. *Circ. Cardiovasc. Genet.* **5,** 647–55 (2012).
9. Holm, H. *et al.* Several common variants modulate heart rate, PR interval and QRS duration. *Nat. Genet.* **42,** 117–22 (2010).
10. Pfeufer, A. *et al.* Common variants at ten loci modulate the QT interval duration in the QTSCD Study. *Nat. Genet.* **41,** 407–414 (2009).
11. Vasan, R. S. *et al.* Genetic Variants Associated With Cardiac Structure and Function. *JAMA* **302,** 168 (2009).
12. Newton-Cheh, C. *et al.* Common variants at ten loci influence QT interval duration in the QTGEN Study. *Nat. Genet.* **41,** 399–406 (2009).
13. Ritchie, M. D. *et al.* Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. *Circulation* **127,** 1377–85 (2013).
14. Nolte, I. M. *et al.* Common genetic variation near the phospholamban gene is associated with cardiac repolarisation: meta-analysis of three genome-wide association studies. *PLoS One* **4,** e6138 (2009).
15. Eijgelsheim, M. *et al.* Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum. Mol. Genet.* **19,** 3885–94 (2010).
16. Verweij, N. *et al.* Genetic Determinants of P Wave Duration and PR Segment. *Circ. Cardiovasc. Genet.* **7,** 475–81 (2014).
17. Pfeufer, A. *et al.* Genome-wide association study of PR interval. *Nat. Genet.* **42,** 153–159 (2010).
18. Sano, M. *et al.* Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations. *Hum. Mol. Genet.* **23,** 6668–76 (2014).
19. Butler, A. A. M. *et al.* Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts. *Circ. Cardiovasc. Genet.* **5,** 639–646 (2012).
20. Hong, K.-W. *et al.* Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. *Hum. Mol. Genet.* **23,** 6659–67 (2014).
21. Sotoodehnia, N. *et al.* Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat. Genet.* **42,** 1068–76 (2010).

22. Smith, J. G. *et al.* Genome-wide association studies of the PR interval in African Americans. *PLoS Genet.* **7,** e1001304 (2011).
23. Petkowski, J. J. *et al.* NRMT2 is an N-terminal monomethylase that primes for its homologue NRMT1. *Biochem. J.* **456,** 453–62 (2013).
24. Bonsignore, L. A. *et al.* NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging. *Mech. Ageing Dev.* **146–148,** 42–52 (2015).
25. Bonsignore, L. A., Butler, J. S., Klinge, C. M. & Schaner Tooley, C. E. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. *Oncotarget* **6,** 12248–12263 (2015).
26. Orr, N. *et al.* A mutation in the atrial-specific myosin light chain gene (MYL4) causes familial atrial fibrillation. *Nat. Commun.* **7,** 11303 (2016).
27. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47,** 435–444 (2015).
28. Shimizu, K. *et al.* SMAP, an Smg GDS-associating protein having arm repeats and phosphorylated by Src tyrosine kinase. *J. Biol. Chem.* **271,** 27013–7 (1996).
29. Shimizu, K., Shirataki, H., Honda, T., Minami, S. & Takai, Y. Complex formation of SMAP/KAP3, a KIF3A/B ATPase motor-associated protein, with a human chromosome-associated polypeptide. *J. Biol. Chem.* **273,** 6591–4 (1998).
30. Hirokawa, N. Stirring up development with the heterotrimeric kinesin KIF3. *Traffic* **1,** 29–34 (2000).
31. Rahmioglu, N. *et al.* Genome-wide enrichment analysis between endometriosis and obesity-related traits reveals novel susceptibility loci. *Hum. Mol. Genet.* **24,** 1185–99 (2015).
32. Landers, J. E. *et al.* Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 9004–9 (2009).
33. Gotoh, M. *et al.* Comprehensive exploration of novel chimeric transcripts in clear cell renal cell carcinomas using whole transcriptome analysis. *Genes. Chromosomes Cancer* **53,** 1018–32 (2014).
34. Telikicherla, D. *et al.* Overexpression of Kinesin Associated Protein 3 (KIFAP3) in Breast Cancer. *J. Proteomics Bioinform.* **5,** 122–126 (2012).
35. Choi, J. *et al.* Kinesin superfamily-associated protein 3 is preferentially expressed in glutamatergic neurons and contributes to the excitatory control of female puberty. *Endocrinology* **149,** 6146–56 (2008).
36. Satoh, A. *et al.* Characterization of human p33/41 (annexin IV), a Ca2+ dependent carbohydrate-binding protein with monoclonal anti-annexin IV antibodies, AS11 and AS17. *Biol. Pharm. Bull.* **20,** 224–9 (1997).
37. Yao, H., Sun, C., Hu, Z. & Wang, W. The role of annexin A4 in cancer. *Front. Biosci. (Landmark Ed.* **21,** 949–57 (2016).
38. Heinick, A. *et al.* Annexin A4 is a novel direct regulator of adenylyl cyclase type 5. *FASEB J.* **29,** fj.14-269837- (2015).
39. Matteo, R. G. & Moravec, C. S. Immunolocalization of annexins IV, V and VI in the failing and non-failing human heart. *Cardiovasc. Res.* **45,** 961–70 (2000).
40. Kimura, T. *et al.* Mouse germ cell-less as an essential component for nuclear integrity. *Mol. Cell. Biol.* **23,** 1304–1315 (2003).
41. Kleiman, S. E. *et al.* Reduced human germ cell-less (HGCL) expression in azoospermic men with severe germinal cell impairment. *J. Androl.* **24,** 670–5
42. Gjerstorff, M. F. *et al.* GAGE cancer-germline antigens are recruited to the nuclear envelope by germ cell-less (GCL). *PLoS One* **7,** e45819 (2012).

43. Fournier, A. *et al.* 1q12 chromosome translocations form aberrant heterochromatic foci associated with changes in nuclear architecture and gene expression in B cell lymphoma. *EMBO Mol. Med.* **2,** 159–71 (2010).
44. Graser, S., Stierhof, Y.-D. & Nigg, E. A. Cep68 and Cep215 (Cdk5rap2) are required for centrosome cohesion. *J. Cell Sci.* **120,** 4321–31 (2007).
45. Man, X., Megraw, T. L. & Lim, Y. P. Cep68 can be regulated by Nek2 and SCF complex. *Eur. J. Cell Biol.* **94,** 162–72 (2015).
46. Kim, J.-H. *et al.* Genome-wide and follow-up studies identify CEP68 gene variants associated with risk of aspirin-intolerant asthma. *PLoS One* **5,** e13818 (2010).
47. Cornejo-García, J. A. *et al.* Variants of CEP68 gene are associated with acute urticaria/angioedema induced by multiple non-steroidal anti-inflammatory drugs. *PLoS One* **9,** e90966 (2014).
48. Bang, M. L. *et al.* The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circ. Res.* **89,** 1065–72 (2001).
49. Gregorio, C. C. *et al.* The NH2 terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity. *J. Cell Biol.* **143,** 1013–27 (1998).
50. Linke, W. A. & Granzier, H. A spring tale: new facts on titin elasticity. *Biophys. J.* **75,** 2613–4 (1998).
51. Siu, B. L. *et al.* Familial dilated cardiomyopathy locus maps to chromosome 2q31. *Circulation* **99,** 1022–6 (1999).
52. Roberts, A. M. *et al.* Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci. Transl. Med.* **7,** 270ra6 (2015).
53. Gerull, B. *et al.* Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat. Genet.* **30,** 201–4 (2002).
54. Herman, D. S. *et al.* Truncations of titin causing dilated cardiomyopathy. *N. Engl. J. Med.* **366,** 619–28 (2012).
55. Akinrinade, O., Alastalo, T.-P. & Koskenvuo, J. W. Relevance of Truncating Titin Mutations in Dilated Cardiomyopathy. *Clin. Genet.* (2016). doi:10.1111/cge.12741
56. Akinrinade, O., Koskenvuo, J. W. & Alastalo, T.-P. Prevalence of Titin Truncating Variants in General Population. *PLoS One* **10,** e0145284 (2015).
57. Marroni, F. *et al.* A genome-wide association scan of RR and QT interval duration in 3 European genetically isolated populations: the EUROSPAN project. *Circ. Cardiovasc. Genet.* **2,** 322–8 (2009).
58. Li, N. *et al.* Ablation of a Ca2+-activated K+ channel (SK2 channel) results in action potential prolongation in atrial myocytes and atrial fibrillation. *J. Physiol.* **587,** 1087–100 (2009).
59. Yu, T. *et al.* Decreased expression of small-conductance Ca2+-activated K+ channels SK1 and SK2 in human chronic atrial fibrillation. *Life Sci.* **90,** 219–227 (2012).
60. Parajuli, N. *et al.* Determinants of ventricular arrhythmias in human explanted hearts with dilated cardiomyopathy. *Eur. J. Clin. Invest.* **45,** 1286–96 (2015).
61. Yu, C.-C. *et al.* Small Conductance Calcium-Activated Potassium Current Is Important in Transmural Repolarization of Failing Human Ventricles. *Circ. Arrhythmia Electrophysiol.* **8,** 667–676 (2015).
62. Terentyev, D. *et al.* Sarcoplasmic reticulum Ca2+ release is both necessary and sufficient for SK channel activation in ventricular myocytes. *AJP Hear. Circ. Physiol.* **306,** H738–H746 (2014).
63. Gui, L. *et al.* Ventricular tachyarrhythmias in rats with acute myocardial infarction involves activation of small-conductance Ca2+-activated K+ channels. *Am. J. Physiol. Heart Circ. Physiol.* **304,** H118-30 (2013).
64. Chang, P.-C. *et al.* Heterogeneous upregulation of apamin-sensitive potassium currents in failing

human ventricles. *J. Am. Heart Assoc.* **2,** e004713 (2013).
65. Mu, Y.-H. *et al.* RyR2 modulates a Ca2+-activated K+ current in mouse cardiac myocytes. *PLoS One* **9,** e94905 (2014).
66. Turker, I. *et al.* Amiodarone inhibits apamin-sensitive potassium currents. *PLoS One* **8,** e70450 (2013).
67. Kim, J.-J. *et al.* Identification of KCNN2 as a susceptibility locus for coronary artery aneurysms in Kawasaki disease using genome-wide association analysis. *J. Hum. Genet.* **58,** 521–5 (2013).
68. Lee, J.-K. *et al.* Consortium-Based Genetic Studies of Kawasaki Disease in Korea: Korean Kawasaki Disease Genetics Consortium. *Korean Circ. J.* **45,** 443–8 (2015).
69. Allen, D. *et al.* SK2 channels are neuroprotective for ischemia-induced neuronal cell death. *J. Cereb. Blood Flow Metab.* **31,** 2302–12 (2011).
70. Orfila, J. E. *et al.* Increasing small conductance Ca^{2+} -activated potassium channel activity reverses ischemia-induced impairment of long-term potentiation. *Eur. J. Neurosci.* **40,** 3179–3188 (2014).
71. McKay, B. M. *et al.* Increasing SK2 channel activity impairs associative learning. *J. Neurophysiol.* **108,** 863–70 (2012).
72. Ohtsuki, G., Piochon, C., Adelman, J. P. P. & Hansel, C. SK2 channel modulation contributes to compartment-specific dendritic plasticity in cerebellar Purkinje cells. *Neuron* **75,** 108–120 (2012).
73. Sun, J. *et al.* UBE3A Regulates Synaptic Plasticity and Learning and Memory by Controlling SK2 Channel Endocytosis. *Cell Rep.* **12,** 449–461 (2015).
74. Willis, M. *et al.* Small-conductance calcium-activated potassium type 2 channels (SK2, KCa2.2) in human brain. *Brain Struct. Funct.* (2016). doi:10.1007/s00429-016-1258-1
75. Cadet, J. L. *et al.* Genome-wide DNA hydroxymethylation identifies potassium channels in the nucleus accumbens as discriminators of methamphetamine addiction and abstinence. *Mol. Psychiatry* (2016). doi:10.1038/mp.2016.48
76. Fakira, A. K., Portugal, G. S., Carusillo, B., Melyan, Z. & Morón, J. A. Increased Small Conductance Calcium-Activated Potassium Type 2 Channel-Mediated Negative Feedback on N-methyl-D-aspartate Receptors Impairs Synaptic Plasticity Following Context-Dependent Sensitization to Morphine. *Biol. Psychiatry* **75,** 105–114 (2014).
77. Tatsuki, F. *et al.* Involvement of Ca2+-Dependent Hyperpolarization in Sleep Duration in Mammals. *Neuron* **90,** 70–85 (2016).
78. Kim, S. H. *et al.* Electrogenic transport and K(+) ion channel expression by the human endolymphatic sac epithelium. *Sci. Rep.* **5,** 18110 (2015).
79. Dolga, A. M. *et al.* Subcellular expression and neuroprotective effects of SK channels in human dopaminergic neurons. *Cell Death Dis.* **5,** e999 (2014).
80. Xiao, Y. *et al.* Overexpression of Trpp5 contributes to cell proliferation and apoptosis probably through involving calcium homeostasis. *Mol. Cell. Biochem.* **339,** 155–61 (2010).
81. Guo, L. *et al.* Identification and characterization of a novel polycystin family member, polycystin-L2, in mouse and human: sequence, expression, alternative splicing, and chromosomal localization. *Genomics* **64,** 241–51 (2000).
82. Volk, T., Schwoerer, A. P., Thiessen, S., Schultz, J.-H. & Ehmke, H. A polycystin-2-like large conductance cation channel in rat left ventricular myocytes. *Cardiovasc. Res.* **58,** 76–88 (2003).
83. Ferdaus, M. Z. & McCormick, J. A. The CUL3/KLHL3-WNK-SPAK/OSR1 pathway as a target for antihypertensive therapy. *Am. J. Physiol. Renal Physiol.* **310,** F1389-96 (2016).
84. Schumacher, F.-R. *et al.* Characterisation of the Cullin-3 mutation that causes a severe form of familial hypertension and hyperkalaemia. *EMBO Mol. Med.* **7,** 1285–306 (2015).
85. Shibata, S., Zhang, J., Puthumana, J., Stone, K. L. & Lifton, R. P. Kelch-like 3 and Cullin 3 regulate electrolyte homeostasis via ubiquitination and degradation of WNK4. *Proc. Natl. Acad. Sci. U. S.*

A. **110,** 7838–43 (2013).
86. Boyden, L. M. *et al.* Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature* **482,** 98–102 (2012).
87. Glover, M. *et al.* Detection of mutations in KLHL3 and CUL3 in families with FHHt (familial hyperkalaemic hypertension or Gordon's syndrome). *Clin. Sci. (Lond).* **126,** 721–6 (2014).
88. Louis-Dit-Picard, H. *et al.* KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. *Nat. Genet.* **44,** 456–60, S1-3 (2012).
89. Saitoh, T. & Katoh, M. Molecular cloning and characterization of human WNT8A. *Int. J. Oncol.* **19,** 123–7 (2001).
90. Cunningham, T. J., Kumar, S., Yamaguchi, T. P. & Duester, G. Wnt8a and Wnt3a cooperate in the axial stem cell niche to promote mammalian body axis extension. *Dev. Dyn.* **244,** 797–807 (2015).
91. Ma, Y. *et al.* The Chromatin Remodeling Protein Bptf Promotes Posterior Neuroectodermal Fate by Enhancing Smad2-Activated wnt8a Expression. *J. Neurosci.* **35,** 8493–506 (2015).
92. Gao, H. *et al.* Polymorphisms and expression of the WNT8A gene in Hirschsprung's disease. *Int. J. Mol. Med.* **32,** 647–52 (2013).
93. Lozano-Velasco, E. *et al.* Pitx2 impairs calcium handling in a dose-dependent manner by modulating Wnt signalling. *Cardiovasc. Res.* **109,** 55–66 (2016).
94. Lai, F. *et al.* cDNA cloning and genomic structure of three genes localized to human chromosome band 5q31 encoding potential nuclear proteins. *Genomics* **70,** 123–30 (2000).
95. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347,** 1260419 (2015).
96. GATT, S. ENZYMIC HYDROLYSIS AND SYNTHESIS OF CERAMIDES. *J. Biol. Chem.* **238,** 3131–3 (1963).
97. Gatt, S. Enzymatic hydrolysis of sphingolipids. I. Hydrolysis and synthesis of ceramides by an enzyme from rat brain. *J. Biol. Chem.* **241,** 3724–30 (1966).
98. Seelan, R. S. *et al.* Human acid ceramidase is overexpressed but not mutated in prostate cancer. *Genes. Chromosomes Cancer* **29,** 137–46 (2000).
99. Norris, J. S. *et al.* Combined therapeutic use of AdGFPFasL and small molecule inhibitors of ceramide metabolism in prostate and head and neck cancers: a status report. *Cancer Gene Ther.* **13,** 1045–51 (2006).
100. Musumarra, G., Barresi, V., Condorelli, D. F. & Scirè, S. A bioinformatic approach to the identification of candidate genes for the development of new cancer diagnostics. *Biol. Chem.* **384,** 321–7 (2003).
101. Saad, A. F. *et al.* The functional effects of acid ceramidase overexpression in prostate cancer progression and resistance to chemotherapy. *Cancer Biol. Ther.* **6,** 1455–60 (2007).
102. Selzner, M. *et al.* Induction of apoptotic cell death and prevention of tumor growth by ceramide analogues in metastatic human colon cancer. *Cancer Res.* **61,** 1233–40 (2001).
103. Beckham, T. H. *et al.* Acid ceramidase-mediated production of sphingosine 1-phosphate promotes prostate cancer invasion through upregulation of cathepsin B. *Int. J. Cancer* **131,** 2034–43 (2012).
104. Kus, G., Kabadere, S., Uyar, R. & Kutlu, H. M. Induction of apoptosis in prostate cancer cells by the novel ceramidase inhibitor ceranib-2. *In Vitro Cell. Dev. Biol. Anim.* **51,** 1056–63 (2015).
105. Zeidan, Y. H. *et al.* Molecular targeting of acid ceramidase: implications to cancer therapy. *Curr. Drug Targets* **9,** 653–61 (2008).
106. Frohbergh, M., He, X. & Schuchman, E. H. The molecular medicine of acid ceramidase. *Biol. Chem.* **396,** 759–65 (2015).
107. Koch, J. *et al.* Molecular cloning and characterization of a full-length complementary DNA encoding human acid ceramidase. Identification Of the first molecular lesion causing Farber

disease. *J. Biol. Chem.* **271,** 33110–5 (1996).
108. Rubboli, G. *et al.* Spinal muscular atrophy associated with progressive myoclonic epilepsy: A rare condition caused by mutations in ASAH1. *Epilepsia* **56,** 692–8 (2015).
109. Huang, Y. *et al.* Elevation of the level and activity of acid ceramidase in Alzheimer's disease brain. *Eur. J. Neurosci.* **20,** 3489–97 (2004).
110. Li, C.-M. M. *et al.* The human acid ceramidase gene (ASAH): structure, chromosomal location, mutation analysis, and expression. *Genomics* **62,** 223–31 (1999).
111. Baranowski, M., Blachnio, A., Zabielski, P. & Gorski, J. Pioglitazone induces de novo ceramide synthesis in the rat heart. *Prostaglandins Other Lipid Mediat.* **83,** 99–111 (2007).
112. Monette, J. S. *et al.* (R)-α-Lipoic acid treatment restores ceramide balance in aging rat cardiac mitochondria. *Pharmacol. Res.* **63,** 23–29 (2011).
113. Wang, L., Lee, K., Malonis, R., Sanchez, I. & Dynlacht, B. D. Tethering of an E3 ligase by PCM1 regulates the abundance of centrosomal KIAA0586/Talpid3 and promotes ciliogenesis. *Elife* **5,** (2016).
114. Zhang, W. *et al.* MiRNA-128 regulates the proliferation and neurogenesis of neural precursors by targeting PCM1 in the developing cortex. *Elife* **5,** (2016).
115. Farina, F. *et al.* The centrosome is an actin-organizing centre. *Nat. Cell Biol.* **18,** 65–75 (2015).
116. Schwaab, J. *et al.* Limited duration of complete remission on ruxolitinib in myeloid neoplasms with PCM1-JAK2 and BCR-JAK2 fusion genes. *Ann. Hematol.* **94,** 233–8 (2015).
117. Sakamoto, S. *et al.* Four polymorphisms of the pericentriolar material 1 (PCM1) gene are not associated with schizophrenia in a Japanese population. *Psychiatry Research* **216,** 288–289 (2014).
118. Stylli, S. S. *et al.* Expression of the adaptor protein Tks5 in human cancer: prognostic potential. *Oncol. Rep.* **32,** 989–1002 (2014).
119. Burger, K. L. *et al.* Src-dependent Tks5 phosphorylation regulates invadopodia-associated invasion in prostate cancer cells. *Prostate* **74,** 134–48 (2014).
120. Blouw, B. *et al.* The invadopodia scaffold protein Tks5 is required for the growth of human breast cancer cells in vitro and in vivo. *PLoS One* **10,** e0121003 (2015).
121. Oikawa, T. *et al.* Tks5-dependent formation of circumferential podosomes/invadopodia mediates cell-cell fusion. *J. Cell Biol.* **197,** 553–68 (2012).
122. Stylli, S. S., I, S. T. T., Kaye, A. H. & Lock, P. *Prognostic significance of Tks5 expression in gliomas*. *Journal of Clinical Neuroscience* **19,** (2012).
123. Wang, F., Chang, J. T.-H., Kao, C. J. & Huang, R. S. High Expression of miR-532-5p, a Tumor Suppressor, Leads to Better Prognosis in Ovarian Cancer Both In Vivo and In Vitro. *Mol. Cancer Ther.* **15,** 1123–31 (2016).
124. Blouw, B., Seals, D. F., Pass, I., Diaz, B. & Courtneidge, S. A. A role for the podosome/invadopodia scaffold protein Tks5 in tumor growth in vivo. *Eur. J. Cell Biol.* **87,** 555–567 (2008).
125. Murphy, D. A. *et al.* A Src-Tks5 pathway is required for neural crest cell migration during embryonic development. *PLoS One* **6,** e22499 (2011).
126. Cejudo-Martin, P. *et al.* Genetic disruption of the sh3pxd2a gene reveals an essential role in mouse development and the existence of a novel isoform of tks5. *PLoS One* **9,** e107674 (2014).
127. Burger, K. L., Davis, A. L., Isom, S., Mishra, N. & Seals, D. F. The podosome marker protein Tks5 regulates macrophage invasive behavior. *Cytoskeleton (Hoboken).* **68,** 694–711 (2011).
128. Vincent, C., Siddiqui, T. A. & Schlichter, L. C. Podosomes in migrating microglia: components and matrix degradation. *J. Neuroinflammation* **9,** 190 (2012).
129. Mesirca, P. *et al.* The G-protein-gated K+ channel, IKACh, is required for regulation of pacemaker activity and recovery of resting heart rate after sympathetic stimulation. *J. Gen. Physiol.* **142,** 113–26 (2013).

130. Mesirca, P. *et al.* Cardiac arrhythmia induced by genetic silencing of 'funny' (f) channels is rescued by GIRK4 inactivation. *Nat. Commun.* **5,** 4664 (2014).
131. Bingen, B. O. *et al.* Atrium-Specific Kir3.x determines inducibility, dynamics, and termination of fibrillation by regulating restitution-driven alternans. *Circulation* **128,** 2732–2744 (2013).
132. Jabbari, J. *et al.* Common polymorphisms in KNCJ5 are associated with early-onset lone atrial fibrillation in Caucasians. *Cardiology* **118,** 116–120 (2011).
133. Wang, F. *et al.* The phenotype characteristics of type 13 long QT syndrome with mutation in KCNJ5 (Kir3.4-G387R). *Hear. Rhythm* **10,** 1500–1506 (2013).
134. Liang, B. *et al.* G-protein-coupled inward rectifier potassium current contributes to ventricular repolarization. *Cardiovasc. Res.* **101,** 175–84 (2014).
135. Molina-Navarro, M. M. *et al.* Differential gene expression of cardiac ion channels in human dilated cardiomyopathy. *PLoS One* **8,** e79792 (2013).
136. Kokunai, Y. *et al.* A Kir3.4 mutation causes Andersen-Tawil syndrome by an inhibitory effect on Kir2.1. *Neurology* **82,** 1058–64 (2014).
137. Azizan, E. A. B. & Brown, M. J. Novel genetic determinants of adrenal aldosterone regulation. *Curr. Opin. Endocrinol. Diabetes. Obes.* **23,** 209–17 (2016).
138. Chen, A. X., Nishimoto, K., Nanba, K. & Rainey, W. E. Potassium channels related to primary aldosteronism: Expression similarities and differences between human and rat adrenals. *Mol. Cell. Endocrinol.* **417,** 141–148 (2015).
139. Gomez, L. *et al.* Association of the KCNJ5 gene with Tourette Syndrome and Attention-Deficit/Hyperactivity Disorder. *Genes. Brain. Behav.* **13,** 535–42 (2014).
140. Han, C., Huang, J. & Waxman, S. G. Sodium channel Nav1.8: Emerging links to human disease. *Neurology* **86,** 473–83 (2016).
141. Yang, T. *et al.* Blocking Scn10a channels in heart reduces late sodium current and is antiarrhythmic. *Circ. Res.* **111,** 322–332 (2012).
142. Verkerk, A. O. *et al.* Functional NaV1.8 channels in intracardiac neurons: The link between SCN10A and cardiac electrophysiology. *Circ. Res.* **111,** 333–343 (2012).
143. Chambers, J. C. *et al.* Genetic variation in SCN10A influences cardiac conduction. *Nat. Genet.* **42,** 149–152 (2010).
144. Denny, J. C. *et al.* Identification of genomic predictors of atrioventricular conduction: using electronic medical records as a tool for genome science. *Circulation* **122,** 2016–21 (2010).
145. Savio-Galimberti, E. *et al.* SCN10A/Nav1.8 modulation of peak and late sodium currents in patients with early onset atrial fibrillation. *Cardiovasc. Res.* **104,** 355–63 (2014).
146. Bezzina, C. R. *et al.* Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death. *Nat. Genet.* **45,** 1044–1049 (2013).
147. Hu, D. *et al.* Mutations in SCN10A are responsible for a large fraction of cases of Brugada syndrome. *J. Am. Coll. Cardiol.* **64,** 66–79 (2014).
148. Park, D. S. & Fishman, G. I. Nav-igating through a complex landscape: SCN10A and cardiac conduction. *J. Clin. Invest.* **124,** 1460–2 (2014).
149. van den Boogaard, M. *et al.* A common genetic variant within SCN10A modulates cardiac SCN5A expression. *J. Clin. Invest.* **124,** 1844–1852 (2014).
150. Aza-Carmona, M. *et al.* NPPB and ACAN, two novel SHOX2 transcription targets implicated in skeletal development. *PLoS One* **9,** e83104 (2014).
151. Liu, C.-F. & Lefebvre, V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Res.* **43,** 8183–203 (2015).
152. Baroti, T. *et al.* Transcription factors Sox5 and Sox6 exert direct and indirect influences on oligodendroglial migration in spinal cord and forebrain. *Glia* **64,** 122–38 (2016).

153. Hersh, C. P. *et al.* SOX5 is a candidate gene for chronic obstructive pulmonary disease susceptibility and is necessary for lung development. *Am. J. Respir. Crit. Care Med.* **183,** 1482–9 (2011).
154. Olesen, M. S. *et al.* Genetic loci on chromosomes 4q25, 7p31, and 12p12 are associated with onset of lone atrial fibrillation before the age of 40 years. *Can. J. Cardiol.* **28,** 191–5 (2012).
155. Della-Morte, D. *et al.* A follow-up study for left ventricular mass on chromosome 12p11 identifies potential candidate genes. *BMC Med. Genet.* **12,** 100 (2011).
156. Wen, Y. *et al.* Integrative analysis of genome-wide association studies and gene expression profiles identified candidate genes for osteoporosis in Kashin-Beck disease patients. *Osteoporos. Int.* **27,** 1041–6 (2016).
157. Jin, J., Chou, C., Lima, M., Zhou, D. & Zhou, X. Systemic Sclerosis is a Complex Disease Associated Mainly with Immune Regulatory and Inflammatory Genes. *Open Rheumatol. J.* **8,** 29–42 (2014).
158. Le Clerc, S. *et al.* Genomewide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genomewide Association Study 03). *J. Infect. Dis.* **200,** 1194–201 (2009).
159. Tu, W. *et al.* Genome-Wide Loci Linked to Non-Obstructive Azoospermia Susceptibility May Be Independent of Reduced Sperm Production in Males with Normozoospermia. *Biol. Reprod.* **92,** 41–41 (2015).
160. Al Zeyadi, M. *et al.* Whole genome microarray analysis in non-small cell lung cancer. *Biotechnol. Biotechnol. Equip.* **29,** 111–118 (2015).
161. Wang, D., Han, S., Wang, X., Peng, R. & Li, X. SOX5 promotes epithelial-mesenchymal transition and cell invasion via regulation of Twist1 in hepatocellular carcinoma. *Med. Oncol.* **32,** 461 (2015).
162. Shiseki, M. *et al.* Identification of the SOX5 gene as a novel IGH-involved translocation partner in BCL2-negative follicular lymphoma with t(12;14)(p12.2;q32). *Int. J. Hematol.* **102,** 633–8 (2015).
163. Kordaß, T. *et al.* SOX5 is involved in balanced MITF regulation in human melanoma cells. *BMC Med. Genomics* **9,** 10 (2016).
164. Kranias, E. G. & Hajjar, R. J. Modulation of cardiac contractility by the phospholamban/SERCA2a regulatome. *Circ. Res.* **110,** 1646–60 (2012).
165. Minamisawa, S. *et al.* Mutation of the phospholamban promoter associated with hypertrophic cardiomyopathy. *Biochem. Biophys. Res. Commun.* **304,** 1–4 (2003).
166. Landstrom, A. P., Adekola, B. A., Bos, J. M., Ommen, S. R. & Ackerman, M. J. PLN-encoded phospholamban mutation in a large cohort of hypertrophic cardiomyopathy cases: summary of the literature and implications for genetic testing. *Am. Heart J.* **161,** 165–71 (2011).
167. Schmitt, J. P. *et al.* Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* **299,** 1410–3 (2003).
168. van Spaendonck-Zwarts, K. Y. *et al.* Genetic analysis in 418 index patients with idiopathic dilated cardiomyopathy: overview of 10 years' experience. *Eur. J. Heart Fail.* **15,** 628–36 (2013).
169. Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat. Genet.* **44,** 670–5 (2012).
170. Harris, T. B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am. J. Epidemiol.* **165,** 1076–87 (2007).
171. Raitoharju, E. *et al.* Common variation in the ADAM8 gene affects serum sADAM8 concentrations and the risk of myocardial infarction in two independent cohorts. *Atherosclerosis* **218,** 127–133 (2011).
172. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am. J. Epidemiol.* **129,** 687–702 (1989).
173. Alonso, A. *et al.* Incidence of atrial fibrillation in whites and African-Americans: the

Atherosclerosis Risk in Communities (ARIC) study. *Am Hear. J* **158,** 111–117 (2009).
174. Conen, D. *et al.* Genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors (GAPP). *Swiss Med. Wkly.* **143,** w13728 (2013).
175. Sinner, M. F. *et al.* Integrating genetic, transcriptional, and functional analyses to identify 5 novel genes for atrial fibrillation. *Circulation* **130,** 1225–35 (2014).
176. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518,** 197–206 (2015).
177. Weeke, P. *et al.* Examining rare and low-frequency genetic variants previously associated with lone or familial forms of atrial fibrillation in an electronic medical record system: a cautionary note. *Circ. Cardiovasc. Genet.* **8,** 58–63 (2015).
178. Fried, L. P. *et al.* The cardiovascular health study: Design and rationale. *Ann. Epidemiol.* **1,** 263–276 (1991).
179. Vaara, S. *et al.* Cohort Profile: the Corogene study. *Int. J. Epidemiol.* **41,** 1265–71 (2012).
180. Dawber, T. R., Meadors, G. F. & Moore, F. E. Epidemiological approaches to heart disease: the Framingham Study. *Am. J. Public Health Nations. Health* **41,** 279–81 (1951).
181. Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am. J. Epidemiol.* **110,** 281–90 (1979).
182. Nieminen, T. *et al.* The Finnish Cardiovascular Study (FINCAVAS): characterising patients with high risk of cardiovascular morbidity and mortality. *BMC Cardiovasc. Disord.* **6,** 9 (2006).
183. Smith, B. H. *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int. J. Epidemiol.* **42,** 689–700 (2013).
184. Schmermund, A. *et al.* Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL Study. *Am. Heart J.* **144,** 212–218 (2002).
185. Winkelmann, B. R. *et al.* Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics* **2,** S1-73 (2001).
186. Smith, J. G., Platonov, P. G., Hedblad, B., Engström, G. & Melander, O. Atrial fibrillation in the Malmö Diet and Cancer study: a study of occurrence, risk factors and diagnostic validity. *Eur. J. Epidemiol.* **25,** 95–102 (2010).
187. Bild, D. E. *et al.* Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am. J. Epidemiol.* **156,** 871–81 (2002).
188. Rasmussen-Torvik, L. J. *et al.* Fasting glucose GWAS candidate region analysis across ethnic groups in the Multiethnic Study of Atherosclerosis (MESA). *Genet. Epidemiol.* **36,** 384–91 (2012).
189. Genes for Cerebral Hemorrhage on Anticoagulation (GOCHA) Collaborative Group. Exploiting common genetic variation to make anticoagulation safer. *Stroke.* **40,** S64-6 (2009).
190. International Stroke Genetics Consortium (ISGC) *et al.* Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat. Genet.* **44,** 328–33 (2012).
191. Lind, L., Fors, N., Hall, J., Marttala, K. & Stenborg, A. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. *Arterioscler. Thromb. Vasc. Biol.* **25,** 2368–75 (2005).
192. Hillege, H. L. *et al.* Urinary albumin excretion predicts cardiovascular and noncardiovascular mortality in general population. *Circulation* **106,** 1777–82 (2002).
193. Shepherd, J. *et al.* Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a

randomised controlled trial. *Lancet* **360,** 1623–30 (2002).
194. Hofman, A. *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur. J. Epidemiol.* **30,** 661–708 (2015).
195. Gioli-Pereira, L. *et al.* Genetic and ElectroNic medIcal records to predict oUtcomeS in Heart Failure patients (GENIUS-HF) - design and rationale. *BMC Cardiovasc. Disord.* **14,** 32 (2014).
196. Völzke, H. *et al.* Cohort profile: the study of health in Pomerania. *Int. J. Epidemiol.* **40,** 294–307 (2011).
197. Hong, Y., Pedersen, N. L., Brismar, K. & de Faire, U. Genetic and environmental architecture of the features of the insulin-resistance syndrome. *Am. J. Hum. Genet.* **60,** 143–52 (1997).
198. Ingelsson, E., Sundström, J., Arnlöv, J., Zethelius, B. & Lind, L. Insulin resistance and risk of congestive heart failure. *JAMA* **294,** 334–41 (2005).
199. Ridker, P. M. *et al.* Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin. Chem.* **54,** 249–55 (2008).
200. Grove, M. L. *et al.* Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One* **8,** e68095 (2013).
201. Holle, R., Happich, M., Löwel, H., Wichmann, H. E. & MONICA/KORA Study Group. KORA--a research platform for population based health research. *Gesundheitswes. (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes* **67 Suppl 1,** S19-25 (2005).
202. Wichmann, H.-E., Gieger, C., Illig, T. & MONICA/KORA Study Group. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1,** S26-30 (2005).
203. Paynter, N. P. *et al.* Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann. Intern. Med.* **150,** 65–72 (2009).
204. Kuriyama, S. *et al.* The Tohoku Medical Megabank Project: Design and Mission. *J. Epidemiol.* **26,** 493–511 (2016).
205. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45,** 580–5 (2013).
206. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–1797 (2012).
207. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40,** D930-4 (2012).
208. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29,** 308–11 (2001).
209. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNiPA: An interactive, genetic variant-centered annotation browser. *Bioinformatics* **31,** 1334–1336 (2015).
210. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).
211. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–75 (2007).
212. Roden, D. M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.* **84,** 362–9 (2008).
213. Wei, W.-Q. *et al.* Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus. *J. Am. Med. Inform. Assoc.* **19,** 219–24
214. Pulley, J., Clayton, E., Bernard, G. R., Roden, D. M. & Masys, D. R. Principles of human subjects protections applied in an opt-out, de-identified biobank. *Clin. Transl. Sci.* **3,** 42–8 (2010).

215. Khurshid, S., Keaney, J., Ellinor, P. T. & Lubitz, S. A. A Simple and Portable Algorithm for Identifying Atrial Fibrillation in the Electronic Medical Record. *Am. J. Cardiol.* **117,** 221–225 (2016).
216. Vermond, R. A. *et al.* Incidence of Atrial Fibrillation and Relationship With Cardiovascular Events, Heart Failure, and Mortality. *J. Am. Coll. Cardiol.* **66,** 1000–1007 (2015).
217. Magnusson, P. K. E. *et al.* The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res. Hum. Genet.* **16,** 317–29 (2013).
218. Hays, J. *et al.* The women's health initiative recruitment methods and results. *Ann. Epidemiol.* **13,** S18–S77 (2003).
219. Anderson, G. L. *et al.* Implementation of the Women's Health Initiative study design. *Ann. Epidemiol.* **13,** S5–S17 (2003).
220. Gerull, B. *et al.* Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. *Nat. Genet.* **36,** 1162–1164 (2004).
221. Cerrone, M. *et al.* Missense mutations in plakophilin-2 cause sodium current deficit and associate with a brugada syndrome phenotype. *Circulation* **129,** 1092–1103 (2014).
222. Peters, S. Arrhythmogenic cardiomyopathy and provocable Brugada ECG in a patient caused by missense mutation in plakophilin-2. *Int. J. Cardiol.* **173,** 317–8 (2014).